

1976

Some sequential inference problems for Polya urns

Pamela Ananis Doctor
Iowa State University

Follow this and additional works at: <https://lib.dr.iastate.edu/rtd>



Part of the [Statistics and Probability Commons](#)

Recommended Citation

Doctor, Pamela Ananis, "Some sequential inference problems for Polya urns " (1976). *Retrospective Theses and Dissertations*. 6203.
<https://lib.dr.iastate.edu/rtd/6203>

This Dissertation is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Retrospective Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.

INFORMATION TO USERS

This material was produced from a microfilm copy of the original document. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the original submitted.

The following explanation of techniques is provided to help you understand markings or patterns which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting thru an image and duplicating adjacent pages to insure you complete continuity.
2. When an image on the film is obliterated with a large round black mark, it is an indication that the photographer suspected that the copy may have moved during exposure and thus cause a blurred image. You will find a good image of the page in the adjacent frame.
3. When a map, drawing or chart, etc., was part of the material being photographed the photographer followed a definite method in "sectioning" the material. It is customary to begin photoing at the upper left hand corner of a large sheet and to continue photoing from left to right in equal sections with a small overlap. If necessary, sectioning is continued again — beginning below the first row and continuing on until complete.
4. The majority of users indicate that the textual content is of greatest value, however, a somewhat higher quality reproduction could be made from "photographs" if essential to the understanding of the dissertation. Silver prints of "photographs" may be ordered at additional charge by writing the Order Department, giving the catalog number, title, author and specific pages you wish reproduced.
5. PLEASE NOTE: Some pages may have indistinct print. Filmed as received.

Xerox University Microfilms

300 North Zeeb Road
Ann Arbor, Michigan 48106

76-28,238

DOCTOR, Pamela Ananis, 1944-
SOME SEQUENTIAL INFERENCE PROBLEMS FOR
POLYA URNS.

Iowa State University, Ph.D., 1976
Statistics

Xerox University Microfilms, Ann Arbor, Michigan 48106

Some sequential inference problems
for Polya urns

by

Pamela Ananis Doctor

A Dissertation Submitted to the
Graduate Faculty in Partial Fulfillment of
The Requirements for the Degree of
DOCTOR OF PHILOSOPHY

Major: Statistics

Approved:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

For the Major Department

Signature was redacted for privacy.

For the Graduate College

Iowa State University
Ames, Iowa

1976

TABLE OF CONTENTS

	Page
I. INTRODUCTION AND REVIEW OF LITERATURE	1
II. RELATIONSHIPS AMONG BERNOULLI, HYPERGEOMETRIC, AND POLYA PROCESSES	9
III. CLOSURE	59
IV. SIMPLICITY, DECOMPOSABILITY, AND THEIR CONSEQUENCES	95
V. COMPLETENESS	105
VI. SEQUENTIAL UMVU ESTIMATION	122
VII. BIBLIOGRAPHY	141
VIII. ACKNOWLEDGMENTS	143

I. INTRODUCTION AND REVIEW OF LITERATURE

In 1946, Girshick, Mosteller, and Savage considered the problem of the uniform-minimum-variance-unbiased (UMVU) estimation of the Bernoulli parameter p for Bernoulli sampling (i.e., sampling with replacement) from a dichotomous population under certain sequential sampling schemes.

Suppose that the dichotomous population is in fact an urn containing W white and B black balls, with $\frac{W}{W+B} = p$. Successive Bernoulli random variables are defined as follows: When the i^{th} ball is drawn from the urn, X_i takes the value 1 (resp., 0) if the drawn ball is white (resp., black), and the ball is returned to the urn. The outcome of n successive draws is represented as the n -vector $\underline{X}_n = (X_1, X_2, \dots, X_n)'$.

When the sampling scheme calls for drawing precisely n_0 balls (single sampling), the UMVU estimate of p , based on \underline{X}_{n_0} , is $\bar{X}_{n_0} = \sum_{i=1}^{n_0} X_i / n_0$; however, \bar{X}_{n_0} is not the UMVU estimate of p for sequential sampling schemes in general.

For any sampling scheme, the outcome of the first n draws can be usefully represented as a "path segment" on the (n, d) -plane, where n is the number of the draw and d is the value of $\sum_{i=1}^n X_i$. The path segment starts at the origin, and at each successive draw, moves one step diagonally (resp., horizontally) when the drawn ball is white ($X_i = 1$) (resp., black ($X_i = 0$)).

It is assumed that one constructs an estimate for p only after sampling has stopped, and so one needs some predetermined plan for deciding when to stop. This is provided by a stop rule S (Lehmann (1949-1950), page 2-34, calls S the sample space) which is a countable collection of path segments which satisfy the following condition: If $\underline{X}_n = (X_1, \dots, X_n) \in S$, there exist no path subsegments $\underline{X}_k = (X_1, \dots, X_k)$, $1 \leq k \leq n-1$, for which $\underline{X}_k \in S$. The term "stop rule" is synonymous with "sequential sampling scheme."

A stop rule partition, call it S_Π , is a partition of the path segments of S into subsets γ of S . The stop rule partitions dealt with by Girshick, Mosteller, and Savage (1946) are defined below.

Definition 1.1 A stop rule partition S_Π is a GMS (Girshick, Mosteller, Savage) stop rule partition if

- (i) $\underline{X}_{n_1} \in \gamma$ and $\underline{X}_{n_2} \in \gamma$, then $n_1 = n_2$ and $\sum_{i=1}^{n_1} X_{1i} = \sum_{i=1}^{n_2} X_{2i}$, and
- (ii) $\underline{X}_n \in \gamma$, then $\sigma(\underline{X}_n) \in \gamma$, provided that $\sigma(\underline{X}_k) \notin S$, for all k , $1 \leq k \leq n-1$, where $\sigma(\underline{X}_n)$ is a permutation of the coordinates of \underline{X}_n .

For a GMS stop rule partition S_Π , there exists, for each set $\gamma \in S_\Pi$, a unique point $\alpha_\gamma = (n_\gamma, d_\gamma)$, where n_γ and d_γ are the values of n and $\sum_{i=1}^n X_i$ common to all path segments in γ . The point α_γ is a stop point; that is, it literally stops every path segment which reaches it, hence the name "boundary", given by the three authors, to the set of points α_γ , $\gamma \in S_\Pi$. Every path segment \underline{X}_n that reaches

the point $\alpha = (n, d)$ has to be a member of the set $\gamma_\alpha \in S_\pi$, so that, as far as the work of Girschick, Mosteller, and Savage (1946) is concerned, the stop points α_γ , $\gamma \in S_\pi$, are synonymous with S_π .

The "stop-points" of the familiar binomial Sequential Probability Ratio Test (SPRT) (See Chapter V of Wald [1947]) in fact form the boundary which corresponds to a GMS stop rule partition.

For a GMS stop rule partition, one may associate with any point α of the boundary, a number $k(\alpha)$, called path segment count (hereafter, "path count") which is the number of path segments of S to reach α .

Girshick, Mosteller, and Savage (1946) exploited the geometric nature of their boundary in the development of a sequential UMVU estimate of p . They identified a sufficient condition for closure of a stop rule (that is, the probability of sampling being terminated equals 1) and developed an unbiased estimate of p defined on the closed boundary which is a function of the path counts. With Wolfowitz (1946) and Savage (1947), they established a necessary and sufficient property for the closed boundary to be boundedly complete.

In a general discussion of sufficiency in the sequential case, Bahadur (1954) seems to indicate the possibility of some of the generalizations of the work of Girshick, Mosteller, and Savage (1946) considered in this thesis, although no details for implementing such generalizations are provided.

In 1956, David and Olkin studied UMVU estimation for GMS stop rule partitions in the case of sampling without replacement from a dichotomous population, and also formulated a certain "generalized

finite population correction".

Sampling without replacement from a dichotomous population (hypergeometric sampling) is also describable in terms of the urn, with the provision that, after each draw, the ball is not returned to the urn. A parameter of interest in this case is also the proportion p of white balls in the urn before sampling begins.

Up to this point we have discussed stop rules for sampling from a dichotomous population with and without replacement. There is another possible replacement policy: replacement with an addition, which we will call a Polya (or urn) replacement policy.

Returning to our model of the urn with white and black balls, the sampling procedure is now as follows: After a white (resp., black) ball is drawn, that ball plus some fixed number β of white (resp., black) balls are added to the urn. Again, the parameter we wish to estimate is p , the original proportion of white balls in the urn. Thus one of the aims of this dissertation is to investigate the UMVU estimate of p under a Polya replacement policy, and we do this in the context of sequential sampling schemes more general than those considered by Girshick, Mosteller, Savage (1946), Wolfowitz (1946), and Savage (1947). We also compare and contrast the properties of stop rules under Bernoulli sampling and Polya (or urn) sampling.

To begin with, in Chapter II, we give a compendium of facts that will be utilized throughout the dissertation. The three types of sampling procedures (Bernoulli, hypergeometric, and Polya) each

give rise to a sequence of random variables $\{X_i\}_{i \in Z}$, where Z is a subset of the positive integers, which defines a stochastic process; the respective processes are the Bernoulli, hypergeometric, and Polya. The probability distributions that underly the three processes are, respectively, the Bernoulli, hypergeometric, and Polya. The relationships among the distributional properties of these three processes are detailed in Section 2.2. In Section 2.3, we introduce a generalization of the finite population correction for a fixed sample size in the Polya sampling context. Section 2.4 establishes conditions under which one can infer the value of sums of Polya probabilities from the known corresponding sum of Bernoulli probabilities, and conversely. As the multinomial distribution is the multivariate version of the Bernoulli distribution, so there exists a multivariate analog of the Polya distribution. All of the relationships between the respective univariate distributions discussed in the chapter also hold for the multivariate case and are given at the end of their respective sections.

In Chapter III, we give a sufficient condition for closure of a stop rule under Polya sampling, similar to the condition given by Wolfowitz (1946) for Bernoulli sampling. We then relate the closure (non-closure) of a stop rule under Bernoulli sampling to closure (non-closure) under Polya sampling. We also show that if a GMS stop rule partition can be approximated for large n by a pair of non-parallel straight lines, say

$$y_1 = a_1 + b_1 n$$

$$y_2 = a_2 + b_2 n, b_1 \neq b_2$$

then the stop rule is not closed with respect to Polya sampling. As a consequence of this, the SPRT with respect to Polya sampling is not closed.

A point that has not been resolved is the discovery of a necessary condition for closure either with respect to Polya sampling or Bernoulli sampling. Also, the extension of the closure condition for stop rules in the context of multivariate Polya and Bernoulli sampling awaits investigation.

In Chapter IV we investigate some geometric properties of boundaries introduced by Girshick, Mosteller, and Savage (1946) and David and Olkin (1956) in a context divorced from their probabilistic origins. We also prove the converse of a theorem by Plackett (1948) for closed, finite boundaries.

In Chapter V, we exploit the geometric properties of boundaries discussed in Chapter IV in developing conditions to ensure the completeness of an arbitrary stop rule partition, with respect to Polya sampling. We then relate the completeness of stop rule partitions with respect to certain classes of statistics under Polya sampling to the same type of completeness under Bernoulli sampling, and conversely.

Chapter VI deals with estimation. In Section 6.2, we establish conditions for the existence of a UMVU estimate of p under Polya

sampling, and show that any finite closed stop rule which satisfies these conditions is a boundary in the sense used by Girshick, Mosteller, and Savage (1946). We compare the conditions for the existence of the UMVU estimate of p in Bernoulli sampling to those for Polya sampling in Section 6.3. Then, in Section 6.4, we introduce a sequential version of the finite population correction for the context of Polya sampling.

Since the random variables $\{X_i\}_{i=1}^{\infty}$ generated by Polya sampling form a stochastic process, our estimation problem can be viewed as one in estimation for a randomly stopped Polya process.

Other work somewhat related to the estimation problem for Polya sampling includes the following three papers.

Audley and Jonckheere (1956), in the context of psychological learning theory, proposed an approximate maximum likelihood method to estimate the parameters for a more complex version of the Polya process than the one we wish to consider here. Bush and Mosteller (1953), again in the context of learning theory, introduced the notion of a linear operator to describe the changes of a stochastic process, and they proposed maximum likelihood estimates for the parameters. However, the Polya process cannot be expressed using linear operators.

One problem with these approaches is that one is left with having to estimate quite a few parameters from only one realization of the process. Both Audley and Jonckheere (1956), and Bush and Mosteller (1953) averaged several realizations, but with mixed

results. We note that in the UMVU approach to the problem, although one only observes one realization of the process, the other possible outcomes play a role in the estimation.

Klotz (1973) proposed fixed-sample-size inference procedures for Bernoulli sampling with a Markov dependence, where both p and the dependence parameter,

$$\lambda = P \{X_i = 1 | X_{i-1} = 1\},$$

are unknown. His approach does not apply to Polya sampling, since the dependence between X_{i-1} and X_i is not Markovian. It is the sum, and not the individual X_i 's, which has the Markov property in this case.

We close the introductory chapter by pointing to possible areas of further research. These include optimal stopping rules for Polya processes similar to the work done by Bahadur (1954) for the general case and by DeGroot (1959) for the binomial case, and the consistency of Polya estimates (see Wolfowitz [1947]) among the class of non-GMS stop rules. Also Bayesian estimation procedures do not seem to have been studied for Polya sampling.

One could investigate the work of Karlin (1974) in fixed-sample-size symmetric sampling schemes for sampling with and without replacement as it applies to Polya single and sequential sampling.

There is the whole field of sequential test procedures, and the work of Robbins (1970) on non-closed tests, to investigate with regard to Polya sampling.

II. RELATIONSHIPS AMONG BERNOULLI, HYPERGEOMETRIC, AND POLYA PROCESSES

2.1. Preliminaries

Consider sampling a population consisting initially of N items of two types. Suppose that, after each sampling occasion, $(1+\beta)$ items of the type sampled are returned to the population. The values of β and N determine the name attached to the stochastic process generated by the sampling procedure, as indicated in the following display:

	N finite	N infinite
$\beta > 0$	Polya	Bernoulli
$\beta = 0$	Bernoulli	Bernoulli
$\beta = -1$	Hypergeometric	Bernoulli

Section 2.2 details the relationships inherent in the display, and the further relationship that Polya processes are beta mixtures of Bernoulli processes. The asymptotic distributions of cumulative counts for Polya and Bernoulli processes also are compared, and extensions are given to the multivariate case of k types.

In Section 2.3, the finite population correction for a fixed sample size is generalized to Polya sampling.

Section 2.4 introduces certain relationships between sums of

terms of type $A_{ij} p^i q^j$ and sums of terms of the type $A_{ij} \frac{W^{[i]} B^{[j]}}{N^{[i+j]}}$, where the A_{ij} 's are constants, $X^{[i]} = X(X+\beta) \dots (X + (i-1)\beta)$, and $N = W + B$. These relationships will be extensively used in Chapters III, V, and VI.

2.2. Distributions and Moments

Polya, Bernoulli, and hypergeometric processes arise from the sampling of a dichotomous population, and, predictably, this similarity is reflected by algebraic similarities in the distributions involved.

Recall the urn containing N balls, W of which are white and B black with $p = \frac{W}{N}$. Suppose that n balls are drawn from this urn, yielding the sample $\{x_1, \dots, x_n\}$, where, as before, x_i is 1 (resp., 0) if the i^{th} ball drawn is white (resp., black).

Sampling the urn with replacement leads to the Bernoulli process under which

$$\begin{aligned}
 P_b(\underline{x}_n) &\triangleq P_b(X_i = x_i; 1 \leq i \leq n) = p^{\sum_{i=1}^n x_i} q^{n - \sum_{i=1}^n x_i} \\
 &= \frac{W^{\sum_{i=1}^n x_i} B^{n - \sum_{i=1}^n x_i}}{N^n}, \quad (2.2.1)
 \end{aligned}$$

where $\underline{x}_n = (x_1, x_2, \dots, x_n)$ and $q = 1-p$. Here successive draws are independent, and sampling can continue indefinitely.

The hypergeometric process, under which

$$P_h(\underline{x}_n) = P_h(X_i = x_i; 1 \leq i \leq n) = \frac{\prod_{i=1}^n x_i - 1 \quad \prod_{j=0}^{n-1} (W-j)}{\prod_{j=0}^{n-1} (N-j)} \prod_{i=1}^n x_i - 1 \quad \prod_{j=0}^{n-1} (B-j), \quad (2.2.2)$$

$\max [0, n-B] \leq \sum_{i=1}^n x_i \leq W$, represents the case of sampling without replacement, where trials are not independent and sampling cannot proceed beyond $n = N$.

In the case of the Polya process, we sample with replacement, plus an addition. That is, after each draw, the drawn ball is replaced, and, in addition, $\beta > 0$ balls of the color drawn are added to the urn, leading to

$$P_u(\underline{x}_n) = P_u(X_i = x_i; 1 \leq i \leq n) = \frac{\prod_{i=1}^n x_i - 1 \quad \prod_{j=0}^{n-1} (W+j\beta)}{\prod_{j=0}^{n-1} (N+j\beta)} \prod_{i=1}^n x_i - 1 \quad \prod_{j=0}^{n-1} (B+j\beta)$$

$$= \frac{W^{\sum_{i=1}^n x_i} B^{\sum_{i=1}^n x_i}}{N^{[n]}}, \quad (2.2.3)$$

where $x^{[i]} = x(x+\beta) \dots (x + (i-1)\beta)$. Here trials are not independent, as in the hypergeometric case, and sampling can continue indefinitely, as in the Bernoulli case.

The relationships among the three processes can be stated succinctly by noting that the Bernoulli and hypergeometric can be considered special cases of the Polya process if we allow and set $\beta = 0$ and -1 respectively. We note as well that all three processes share the property of exchangeability; that is all k -dimensional marginal distributions are the same (see page 228 of Feller, 1971). In the case of the Bernoulli process (when $\beta = 0$), this reduces to independence.

The Bernoulli process can also be related to the Polya and hypergeometric processes in another way. Suppose for example that $\beta > 0$ is fixed and that the urn initially contains, respectively, $[m\theta]$ and $[m\rho]$ white and black balls, where $[y]$ represents the largest integer in y , and $0 \leq \rho = 1 - \theta \leq 1$. Then

$$P_u(\underline{x}_n) = \frac{([m\theta])^{[\sum_{i=1}^n x_i]} ([m\rho])^{[n - \sum_{i=1}^n x_i]}}{([m\theta] + [m\rho])^{[n]}} \xrightarrow{m \rightarrow \infty} \theta^{\sum_{i=1}^n x_i} \rho^{n - \sum_{i=1}^n x_i} \\ = P_b(\underline{x}_n). \quad (2.2.4)$$

In other words, the Bernoulli process is, in the sense of (2.2.4), a limiting form of a Polya process.

In addition, the Bernoulli and Polya processes are related by the fact that the Polya density is a beta mixture of binomials (see Johnson and Kotz (1969), page 229):

$$P_u(x_n) = \int_0^1 P_b(x_n) \frac{\Gamma(\frac{N}{\beta})}{\Gamma(\frac{W}{\beta}) \Gamma(\frac{B}{\beta})} p^{\frac{W}{\beta}-1} q^{\frac{B}{\beta}-1} dp. \quad (2.2.5)$$

Turning now to the first and second moments of $\sum_{i=1}^n X_i$ under the three processes, exchangeability leads, for all three, to

$$E\left(\sum_{i=1}^n X_i\right) = np, \quad (2.2.6)$$

where $p = \frac{W}{N}$, and $n \leq N$ in the hypergeometric case.

The variances of $\sum_{i=1}^n X_i$ under the three processes differ. The Bernoulli variance is

$$\text{Var}_b\left(\sum_{i=1}^n X_i\right) = npq; \quad (2.2.7)$$

the hypergeometric variance is

$$\text{Var}_h\left(\sum_{i=1}^n X_i\right) = npq \left(\frac{N-n}{N-1}\right), \quad (2.2.8)$$

where the term $\frac{N-n}{N-1}$ is called the finite population correction; and the Polya variance is

$$\text{Var}_u\left(\sum_{i=1}^n X_i\right) = npq \left(\frac{N+n\beta}{N+\beta}\right). \quad (2.2.9)$$

Note that the factor $\frac{N + n\beta}{N + \beta}$ is entirely analogous to the hypergeometric finite population correction.

Note as well that the hypergeometric process assigns a smaller variance, and the Polya process a larger variance, than the Bernoulli, to $\sum_{i=1}^n X_i$, which seems in keeping with the underlying replacement policies.

Applying the Central Limit Theorem to the Bernoulli process,

$$\frac{\sum_{i=1}^n X_i - np}{\sqrt{npq}} \xrightarrow[n \rightarrow \infty]{\mathcal{L}} N(0,1). \quad (2.2.10)$$

On the other hand, Freedman (1965) found that, for the Polya process,

$$\frac{\sum_{i=1}^n X_i}{n} \xrightarrow[n \rightarrow \infty]{a.s.} Z_{\frac{W}{\beta}, \frac{B}{\beta}}, \quad (2.2.11)$$

where $Z_{\frac{W}{\beta}, \frac{B}{\beta}}$ denotes a random variable distributed according to the beta density, $\mathcal{B}_{\frac{W}{\beta}, \frac{B}{\beta}}$, with parameters $\frac{W}{\beta}$ and $\frac{B}{\beta}$, which appears in (2.2.5), so that, in particular,

$$\frac{\sum_{i=1}^n X_i}{n} \xrightarrow[n \rightarrow \infty]{\mathcal{L}} Z_{\frac{W}{\beta}, \frac{B}{\beta}}.$$

The dependence on n of the variances in (2.2.7) and (2.2.9) are reflected in the norming constants of (2.2.10) and (2.2.11), and also in the fact, as will be seen in Chapter III, that more stop rules are closed under Polya sampling than under Bernoulli sampling.

The k -color versions of the above distributional facts can be readily obtained, and are given below.

Consider an urn initially containing W_1 balls of color 1, W_2 balls of color 2, ..., and W_k balls of color k . Let $N = \sum_{i=1}^k W_i$, $p_1 = \frac{W_1}{N}$, $p_2 = \frac{W_2}{N}$, ..., and $p_k = \frac{W_k}{N} = 1 - \sum_{i=1}^{k-1} p_i$. When a ball is drawn from the urn, the result is usefully indicated by the vector $(x_1, x_2, \dots, x_{k-1})'$ where $x_1 (x_2, \dots, x_{k-1}) = 1$ or 0 depending on whether or not a ball of color 1 (2, ..., $k-1$) is drawn, and, if n balls are drawn from the urn, the outcome is then indicated

$$\text{by the matrix } \underline{x}_n = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \vdots & \vdots & & \vdots \\ x_{k-11} & x_{k-12} & & x_{k-1n} \end{bmatrix}.$$

The random vector taking values $(x_1, \dots, x_{k-1})'$ will be denoted by $(X_1, \dots, X_{k-1})'$. With this notation, $\sum_{i=1}^n x_{1i}$, $\sum_{i=1}^n x_{2i}$, ...,

$\sum_{j=1}^{k-1} \sum_{i=1}^n x_{ji}$ respectively count the numbers of balls of colors

1, 2, ..., and k obtained in n draws.

In the case of sampling with replacement, the analogue of (2.2.1) is multinomial:

$$P_b(\underline{x}_n) \triangleq \prod_{j=1}^{k-1} p_j^{\sum_{i=1}^n x_{ji}} p_k^{(n - \sum_{j=1}^{k-1} \sum_{i=1}^n x_{ji})}$$

$$= \frac{\prod_{j=1}^{k-1} w_j^{\sum_{i=1}^n x_{ji}} w_k^{n - \sum_{j=1}^{k-1} \sum_{i=1}^n x_{ji}}}{N^n}; \quad (2.2.1a)$$

the analogue of (2.2.2) is

$$P_h(\underline{x}_n) \triangleq \frac{\prod_{\ell=1}^{k-1} \left(\prod_{j=0}^{\sum_{i=1}^n x_{\ell i} - 1} (w_{\ell} - j) \right) \prod_{j=0}^{n - \sum_{\ell=1}^{k-1} \sum_{i=1}^n x_{\ell i} - 1} (w_k - j)}{\prod_{j=0}^{n-1} (N - j)}, \quad (2.2.2a)$$

where $0 \leq \sum_{i=1}^n x_{1i} \leq w_1$, $0 \leq \sum_{i=1}^n x_{2i} \leq w_2$, ..., and

$0 \leq n - \sum_{\ell=1}^{k-1} \sum_{i=1}^n x_{\ell i} \leq w_k$; and the analogue of (2.2.3), adding $\beta > 0$

balls of the color drawn, is

$$\begin{aligned}
P_u(\underline{x}_n) &\triangleq \frac{\prod_{\ell=1}^{k-1} \left(\prod_{j=0}^n x_{\ell j}^{-1} \right) \prod_{j=0}^{n - \sum_{\ell=1}^{k-1} \sum_{i=1}^n x_{\ell i}^{-1}} (w_{\ell} + j\beta)}{\prod_{j=0}^{n-1} (N + j\beta)} \\
&= \frac{\prod_{\ell=1}^{k-1} \left(w_{\ell}^{\left[\sum_{i=1}^n x_{\ell i} \right]} \right) w_k^{\left[n - \sum_{\ell=1}^{k-1} \sum_{i=1}^n x_{\ell i} \right]}}{N^{[n]}}. \quad (2.2.3a)
\end{aligned}$$

The three multivariate processes are also exchangeable, and, if we allow and set $\beta = 0$ and -1 , the multivariate Polya process becomes, respectively, the multinomial and multivariate hypergeometric process. The limiting relationship (2.2.4) extends trivially as well.

It can also be shown that a multivariate Polya process is a Dirichlet mixture of multinomial processes:

$$\begin{aligned}
P_u(\underline{x}_n) &= \int P_b(\underline{x}_n) \frac{\Gamma\left(\frac{N}{\beta}\right)}{\prod_{i=1}^k \Gamma\left(\frac{w_i}{\beta}\right)} \prod_{\ell=1}^{k-1} p_{\ell}^{\frac{w_{\ell}}{\beta}-1} dp_1 dp_2 \dots dp_k \\
&\quad p_1, p_2, \dots, p_k \geq 0 \\
&\quad \text{and } \sum_{i=1}^k p_i = 1 \\
&\equiv \int_{\substack{p_1, p_2, \dots, p_k \geq 0 \\ \sum_{i=1}^k p_i = 1}} P_b(\underline{x}_n) \mathcal{D}_{\frac{w_1}{\beta}, \dots, \frac{w_k}{\beta}}(p_1, \dots, p_k) dp_1 \dots dp_k. \quad (2.2.5a)
\end{aligned}$$

The expectations and variances are as in the univariate case; the covariances are given by

$$\text{Cov}_b \left(\sum_{i=1}^n X_{li}, \sum_{i=1}^n X_{mi} \right) = \frac{-nW_l W_m}{N^2},$$

$$\text{Cov}_h \left(\sum_{i=1}^n X_{li}, \sum_{i=1}^n X_{mi} \right) = - \frac{nW_l W_m}{N^2} \left[\frac{N-n}{N-1} \right],$$

and

$$\text{Cov}_u \left(\sum_{i=1}^n X_{li}, \sum_{i=1}^n X_{mi} \right) = - \frac{nW_l W_m}{N^2} \left[\frac{N+n\beta}{N+\beta} \right] \quad (2.2.12)$$

(see page 301 of Johnson and Kotz [1969]). Hence, in view of (2.2.8) and (2.2.9), the variance-covariance matrix for the hypergeometric or Polya process is given by the product of the multinomial variance-covariance matrix and the factor $\frac{N-n}{N-1}$ or $\frac{N+n\beta}{N+\beta}$, respectively.

The analogue of (2.2.10) for $k = 3$ is now

$$\frac{\left(\sum_{i=1}^n X_{1i}, \sum_{i=1}^n X_{2i} \right) - n(p_1, p_2)}{\sqrt{n}} \xrightarrow[n \rightarrow \infty]{\mathcal{L}} N((0,0), \Sigma), \quad (2.2.10a)$$

$$\text{where } \Sigma = \begin{bmatrix} p_1(1-p_1) & -p_1p_2 \\ -p_1p_2 & p_2(1-p_2) \end{bmatrix};$$

This is given by Theorem 2.4.3, page 74, of Anderson (1958).

The multivariate analog of (2.2.11) is given in the following:

Lemma 2.2.1 If the distribution of $(\sum_{i=1}^n X_{1i}, \dots, \sum_{i=1}^n X_{k-1i})$ is given by (2.2.3a), then

$$\left(\frac{\sum_{i=1}^n X_{1i}}{n}, \dots, \frac{\sum_{i=1}^n X_{k-1i}}{n} \right) \xrightarrow[n \rightarrow \infty]{a.s.} Z_{\frac{w_1}{\beta}, \dots, \frac{w_k}{\beta}}, \quad (2.2.11a)$$

where $Z_{\frac{w_1}{\beta}, \dots, \frac{w_k}{\beta}}$ is a Dirichlet random variable with parameters

$$\frac{w_1}{\beta}, \frac{w_2}{\beta}, \dots, \text{ and } \frac{w_k}{\beta}.$$

Proof: For $t_1, \dots, t_{k-1} \in [0,1]$, with $\sum_{i=1}^{k-1} t_i \leq 1$,

$$\lim_{n \rightarrow \infty} P_u \left(\sum_{i=1}^n X_{1i} \leq nt_1, \dots, \sum_{i=1}^n X_{k-1i} \leq nt_{k-1}; w_1, \dots, w_k, \beta \right) =$$

$$\lim_{n \rightarrow \infty} \int_{\substack{p_1, \dots, p_k \geq 0 \\ \sum_{i=1}^k p_i = 1}} p_b \left(\sum_{i=1}^n X_{1i} \leq nt_1, \dots, \sum_{i=1}^n X_{k-1i} \leq nt_{k-1}; p_1, \dots, p_k \right)$$

$$\cdot \mathcal{D}_{\frac{w_1}{\beta}, \dots, \frac{w_k}{\beta}} (p_1, \dots, p_k) dp_1 \dots dp_k$$

by (2.25a), which, by the Lebesgue Dominated Convergence Theorem, equals

$$\int_{\substack{p_1, \dots, p_k \geq 0 \\ \exists \sum_{i=1}^k p_i = 1}} \lim_{n \rightarrow \infty} p_n \left(\sum_{i=1}^n X_{1i} \leq nt_1, \dots, \sum_{i=1}^n X_{k-1,i} \leq nt_{k-1}; p_1, \dots, p_k \right) \\ \cdot \mathcal{D}_{\frac{w_1}{\beta}, \dots, \frac{w_k}{\beta}}(p_1, \dots, p_k) dp_1 \dots dp_k,$$

which, in view of (2.2.9a), equals

$$\int_{\substack{p_1, \dots, p_k \geq 0 \\ \exists \sum_{i=1}^k p_i = 1}} \delta(t_1, \dots, t_{k-1}; p_1, \dots, p_k) \mathcal{D}_{\frac{w_1}{\beta}, \dots, \frac{w_k}{\beta}}(p_1, \dots, p_k) dp_1 \dots dp_k,$$

where $\delta(t_1, \dots, t_{k-1}; p_1, \dots, p_k) = \begin{cases} 1 & \text{if } t_i \geq p_i, i=1, \dots, k-1 \\ 0 & \text{otherwise,} \end{cases}$

$$= \int_{\substack{p_1, \dots, p_k \geq 0 \\ \exists p_i \leq t_i, i=1, \dots, k-1 \\ \text{and } \sum_{i=1}^k p_i = 1}} \mathcal{D}_{\frac{w_1}{\beta}, \dots, \frac{w_k}{\beta}}(p_1, \dots, p_k) dp_1 \dots dp_k.$$

Therefore

$$\left(\frac{\sum_{i=1}^n X_{1i}}{n}, \dots, \frac{\sum_{i=1}^n X_{k-1 i}}{n} \right) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \frac{Z_{W_1}}{\beta}, \dots, \frac{Z_{W_k}}{\beta}. \quad (2.2.13)$$

It can be easily shown that for $j=1, \dots, k-1$,

$$\frac{W_j + \sum_{i=1}^n X_{ji} \beta}{N + n\beta}$$

is a bounded martingale and hence has an almost-sure limit, Y_j , by the Martingale Convergence Theorem. Notice that

$$\lim_{n \rightarrow \infty} \left(\frac{\frac{W_j + \sum_{i=1}^n X_{ji} \beta}{n}}{\frac{N}{n} + \frac{n\beta}{n}} - \frac{\sum_{i=1}^n X_{ji}}{n} \right) = 0,$$

pointwise for all positive integers W , B , and β ; therefore,

$$\frac{\sum_{i=1}^n X_{ji}}{n} \xrightarrow{\text{a.s.}} Y_j$$

and

$$\left(\frac{\sum_{i=1}^n X_{1i}}{n}, \dots, \frac{\sum_{i=1}^n X_{k-1 i}}{n} \right) \xrightarrow{\text{a.s.}} (Y_1, \dots, Y_{k-1}). \quad (2.2.14)$$

Since almost sure convergence implies convergence in distribution, it follows from (2.2.13) and (2.2.14) that (Y_1, \dots, Y_{k-1}) has the indicated Dirichlet distribution.

qed

It should be noted that when $k=2$, the above is an alternate proof of (2.2.11) to the one given by Freedman (1965).

2.3. Polya Sampling of Finite Populations and a Polya Sampling Correction

We saw in Section 2.2 that the variance of cumulative counts for the hypergeometric or Polya process is the product of the binomial variance npq and a certain factor $(\frac{N-n}{N-1}$ or $\frac{N+n\beta}{N+\beta}$, respectively) which, in the hypergeometric case, is known as the finite population correction.

Recall now that this finite population correction $\frac{N-n}{N-1}$ actually applies not only to hypergeometric processes, but to finite population sampling generally, where it takes the form $\text{Var}_h \left(\sum_{i=1}^n X_i \right) = n \sigma_x^2 \left(\frac{N-n}{N-1} \right)$, where σ_x^2 is the variance of the finite population sampled.

It is now natural to ask whether the Polya factor $\frac{N+n\beta}{N+\beta}$ also applies as well when an arbitrary finite population is "Polya-sampled," i.e., is successively sampled with replacement, on each sampling occasion, of $(\beta + 1)$ copies of the item drawn. The question is answered in the affirmative in the following:

Lemma 2.3.1 Suppose a sample $\{X_1, \dots, X_n\}$ is drawn from a population of N numbers z_1, z_2, \dots, z_N , in such a way that, after each draw, the number drawn and $\beta > 0$ copies of it are added to the population. Then the variance of the sum $\sum_{i=1}^n X_i$ is $n \sigma_z^2 \frac{N+n\beta}{N+\beta}$, where $\sigma_z^2 = \frac{1}{N} \sum_{i=1}^N (z_i - \bar{z})^2$.

Proof: Let n_j be the number of times that z_j is drawn in n trials. Then $\sum_{i=1}^n X_i$ equals $\sum_{j=1}^N n_j z_j$, and

$$\begin{aligned}
\text{Var}_u \left(\sum_{i=1}^n X_i \right) &= \text{Var}_u \left(\sum_{j=1}^N n_j z_j \right) \\
&= \sum_{j=1}^N z_j^2 \text{Var}_u(n_j) + 2 \sum_{j=1}^N \sum_{k=j+1}^N z_j z_k \text{Cov}_u(n_j, n_k), \quad (2.3.1)
\end{aligned}$$

which, substituting (2.2.9) and (2.2.12) in (2.3.1), where $\frac{w_j}{N} = \frac{1}{N}$, becomes

$$\begin{aligned}
&n \frac{N+n\beta}{N+\beta} \cdot \frac{N-1}{N^2} \sum_{j=1}^N z_j^2 - 2n \frac{N+n\beta}{N+\beta} \frac{1}{N^2} \sum_{j=1}^N \sum_{k=j+1}^N z_j z_k \\
&= n \frac{N+n\beta}{N+\beta} \left[\frac{1}{N} \sum_{j=1}^N z_j^2 - \frac{1}{N^2} \sum_{j=1}^N \sum_{k=1}^N z_j z_k \right] \\
&= n \sigma_z^2 \frac{N+n\beta}{N+\beta}.
\end{aligned}$$

qed

2.4. The Relationships Between "Polya" Identities and "Bernoulli" Identities

This section is devoted to the exploration of the relationship between identities in series of "Polya-like" terms, $\frac{W[i]_B[j]}{N[i+j]}$, and the corresponding series in "Bernoulli-like" terms, $p^i q^j$.

In particular, in Theorems 2.4.1 and 2.4.2, we ask under what conditions

$$\sum_{(i,j)} A_{ij} p^i q^j = \sum_{(i,j)} C_{ij} p^i q^j \quad (2.4.1)$$

implies

$$\sum_{(i,j)} A_{ij} \frac{W[i]_B[j]}{N[i+j]} = \sum_{(i,j)} C_{ij} \frac{W[i]_B[j]}{N[i+j]}. \quad (2.4.2)$$

Two versions of the converse are studied in Theorems 2.4.3 and 2.4.4.

Theorem 2.4.1 sets conditions on the convergence of

$$\sum_{(i,j)} A_{ij} \frac{W[i]_B[j]}{N[i+j]}$$

and

$$\sum_{(i,j)} C_{ij} \frac{W[i]_B[j]}{N[i+j]}$$

to ensure the applicability of Fubini's Theorem to (2.2.5), and requires that (2.4.1) hold for all $p \in (0,1)$. Its corollary details

the situation when the A_{ij} 's and C_{ij} 's are bounded in absolute value.

Theorem 2.4.1 Consider two doubly indexed sequences $\{A_{ij}\}$ and $\{C_{ij}\}$, and three positive integers W , B , and β . Suppose that

(i) either

$$\sum_{(i,j)} |A_{ij}| \frac{W [i] B [j]}{N [i+j]} < +\infty \quad (2.4.3)$$

or

$$\int_0^1 \left[\sum_{(i,j)} |A_{ij}| p^i q^j \right] \mathcal{B}_{\frac{W}{\beta}, \frac{B}{\beta}}(p) dp < +\infty, \quad (2.4.4)$$

(ii) either

$$\sum_{(i,j)} |C_{ij}| \frac{W [i] B [j]}{N [i+j]} < +\infty \quad (2.4.5)$$

or

$$\int_0^1 \left[\sum_{(i,j)} |C_{ij}| p^i q^j \right] \mathcal{B}_{\frac{W}{\beta}, \frac{B}{\beta}}(p) dp < +\infty, \quad (2.4.6)$$

and, for all $p \in (0,1)$,

$$\sum_{(i,j)} A_{ij} p^i q^j = \sum_{(i,j)} C_{ij} p^i q^j. \quad (2.4.7)$$

Then

$$\sum_{(i,j)} A_{ij} \frac{w[i]_B[j]}{N[i+j]} = \sum_{(i,j)} C_{ij} \frac{w[i]_B[j]}{N[i+j]} .$$

Proof: Let I denote the set of pairs (i,j) of non-negative integers, and let \mathcal{Q} denote the measure on I assigning unit weight to each pair. Define $\overline{\mathcal{B}}_{\frac{w}{\beta}, \frac{B}{\beta}}$ to be the measure corresponding to $\mathcal{B}_{\frac{w}{\beta}, \frac{B}{\beta}}$. It is clear that (2.4.4) ensures the integrability of $A_{ij} p^i q^j$ over $I \times (0,1)$, with respect to the product measure $\mathcal{Q} \times \overline{\mathcal{B}}_{\frac{w}{\beta}, \frac{B}{\beta}}$, so by Fubini's Theorem (Theorem 19, page 269 of Royden [1968]),

$$\begin{aligned} & \int_0^1 \sum_{(i,j)} A_{ij} p^i q^j \mathcal{B}_{\frac{w}{\beta}, \frac{B}{\beta}}(p) dp \\ &= \sum_{(i,j)} \int_0^1 A_{ij} p^i q^j \mathcal{B}_{\frac{w}{\beta}, \frac{B}{\beta}}(p) dp. \end{aligned} \tag{2.4.8}$$

On the other hand, (2.4.8) holds as well under (2.4.3), since in view of (2.2.5),

$$\sum_{(i,j)} |A_{ij}| \frac{w[i]_B[j]}{N[i+j]} = \sum_{(i,j)} \int_0^1 |A_{ij}| p^i q^j \mathcal{B}_{\frac{w}{\beta}, \frac{B}{\beta}}(p) dp,$$

so that $A_{ij} p^i q^j$ also is integrable over $I \times (0,1)$ with respect to $\mathcal{Q} \times \mathcal{B}_{\frac{w}{\beta}, \frac{B}{\beta}}$.

Similar considerations hold for $C_{ij} p^i q^j$, so that by (2.2.5), (2.4.8), (2.4.7), and the C_{ij} analogues of (2.4.8) and (2.2.5),

$$\begin{aligned}
 \sum_{(i,j)} A_{ij} \frac{W [i] B [j]}{N [i+j]} &= \sum_{(i,j)} \int_0^1 A_{ij} p^i q^j \mathcal{B}_{\frac{W}{\beta}, \frac{B}{\beta}}(p) dp \\
 &= \int_0^1 \sum_{(i,j)} A_{ij} p^i q^j \mathcal{B}_{\frac{W}{\beta}, \frac{B}{\beta}}(p) dp \\
 &= \int_0^1 \sum_{(i,j)} C_{ij} p^i q^j \mathcal{B}_{\frac{W}{\beta}, \frac{B}{\beta}}(p) dp \\
 &= \sum_{(i,j)} \int_0^1 C_{ij} p^i q^j \mathcal{B}_{\frac{W}{\beta}, \frac{B}{\beta}}(p) dp \\
 &= \sum_{(i,j)} C_{ij} \frac{W [i] B [j]}{N [i+j]}.
 \end{aligned}$$

qed

Corollary 2.4.1 Consider two doubly indexed sequences $\{A_{ij}\}$ and $\{C_{ij}\}$ and three positive integers W, B and β . Suppose that $W, B > \beta$, and $\sup_{(i,j)} |A_{ij}|, \sup_{(i,j)} |C_{ij}| < K < \infty$. If

$$\sum_{(i,j)} A_{ij} p^i q^j = \sum_{(i,j)} C_{ij} p^i q^j$$

for almost all $p \in (0,1)$, then

$$\sum_{(i,j)} A_{ij} \frac{W^{[i]} B^{[j]}}{N^{[i+j]}} = \sum_{(i,j)} C_{ij} \frac{W^{[i]} B^{[j]}}{N^{[i+j]}}.$$

Proof: We need only show (2.4.4) and (2.4.6) hold to satisfy the conditions of the theorem. So

$$\begin{aligned} \int_0^1 \left[\sum_{(i,j)} |A_{ij}| p^i q^j \right] \mathcal{B}_{\frac{W}{\beta}, \frac{B}{\beta}}(p) dp &\leq \int_0^1 \frac{1}{pq} \mathcal{B}_{\frac{W}{\beta}, \frac{B}{\beta}}(p) dp \\ &= K \frac{(N-\beta)(N-2\beta)}{(W-\beta)(B-\beta)} < +\infty, \end{aligned}$$

and similarly for the C_{ij} 's.

qed

It seems reasonable that inequalities, as well as equalities, are preserved under the mixing operation given in (2.2.5).

Lemma 2.4.1 Consider two doubly-indexed sequences $\{A_{ij}\}$ and $\{C_{ij}\}$ which satisfy conditions (i) and (ii) of Theorem 2.4.4. If, for all $p \in (0,1)$,

$$\sum_{(i,j)} A_{ij} p^i q^j \leq \sum_{(i,j)} C_{ij} p^i q^j,$$

then

$$\sum_{(i,j)} A_{ij} \frac{W^{[i]} B^{[j]}}{N^{[i+j]}} \leq \sum_{(i,j)} C_{ij} \frac{W^{[i]} B^{[j]}}{N^{[i+j]}}$$

for all positive integers W , B , and β .

Proof: Fix arbitrary W , B , and β . In view of conditions (i) and (ii), the following integration is possible:

$$\int_0^1 \sum_{(i,j)} A_{ij} p^i q^j \mathcal{B}_{\frac{W}{\beta}, \frac{B}{\beta}}(p) dp \leq$$

$$\int_0^1 \sum_{(i,j)} C_{ij} p^i q^j \mathcal{B}_{\frac{W}{\beta}, \frac{B}{\beta}}(p) dp.$$

Applying Fubini's Theorem (see Theorem 19, page 269 of Royden [1968]), we have

$$\sum_{(i,j)} A_{ij} \int_0^1 p^i q^j \mathcal{B}_{\frac{W}{\beta}, \frac{B}{\beta}}(p) dp \leq$$

$$\sum_{(i,j)} C_{ij} \int_0^1 p^i q^j \mathcal{B}_{\frac{W}{\beta}, \frac{B}{\beta}}(p) dp,$$

which is equivalent, by (2.2.5), to

$$\sum_{(i,j)} A_{ij} \frac{W^{[i]} B^{[j]}}{N^{[i+j]}} \leq \sum_{(i,j)} C_{ij} \frac{W^{[i]} B^{[j]}}{N^{[i+j]}}.$$

qed

Theorem 2.4.2 also requires absolute summability of

$$\sum_{(i,j)} A_{ij} \frac{W^{[i]} B^{[j]}}{N^{[i+j]}}$$

and

$$\sum_{(i,j)} C_{ij} \frac{W^{[i]} B^{[j]}}{N^{[i+j]}} ,$$

but approaches the problem from the theory of power series; so it places constraints on the difference $|A_{ij} - C_{ij}|$ to ensure the absolute convergence of $\sum_{i=0}^{\infty} \sum_{j=0}^{\infty} (A_{ij} - C_{ij}) p^i q^j$, when the binomial expansion is substituted for q^j . But in contrast to Theorem 2.4.1, (2.4.1) need only hold on a subset $(0,b)$ of $(0,1)$.

The proof of Theorem 2.4.2 requires a result for sums of "Polya-type" terms analogous to the binomial expansion. It is given in

Lemma 2.4.2 For all positive integers i, N, k, r, β , and for $p \in [0,1]$,

$$\sum_{\ell=0}^k (-1)^\ell \binom{k}{\ell} (N + (i+r+\ell)\beta)^{[k-\ell]} (Np+i\beta)^{[\ell]} = (Nq + r\beta)^{[k]} \quad (2.4.9)$$

Proof: Let i, N, β be fixed, but arbitrary integers and $p \in [0,1]$.

The proof will be by induction on k and r .

Let $k = 1$; r is arbitrary.

$$\sum_{\ell=0}^1 (-1)^\ell \binom{1}{\ell} (N + (i+r+\ell)\beta)^{[1-\ell]} (Np+i\beta)^{[\ell]}$$

$$= (N + (i+r)\beta) - (Np + i\beta) = Nq + r\beta.$$

Assume it is true for $k = m-1$ and all $r \geq 0$. Let $k = m$, $r = 0$, then

$$\begin{aligned} & \sum_{\ell=0}^m (-1)^\ell \binom{m}{\ell} (N + (i+\ell)\beta)^{[m-\ell]} (Np + i\beta)^{[\ell]} \\ &= (N + i\beta)^{[m]} - \binom{m}{1} (N + (i+1)\beta)^{[m-1]} (Np + i\beta) \\ & \quad + \binom{m}{2} (N + (i+2)\beta)^{[m-2]} (Np + i\beta)^{[2]} \\ & \quad + \dots + (-1)^{m-2} \binom{m}{m-2} (N + (i+m-2)\beta)^{[2]} (Np+i\beta)^{[m-2]} \\ & \quad + (-1)^{m-1} \binom{m}{m-1} (N + (i+m-1)\beta) (Np + i\beta)^{[m-1]} \\ & \quad + (-1)^m (Np + i\beta)^{[m]}, \end{aligned}$$

and using the fact that $\binom{k}{i} = \binom{k-1}{i-1} + \binom{k-1}{i}$, $i=1, \dots, k-1$, we have

$$\begin{aligned} & (N+i\beta)^{[m]} - \left[\binom{m-1}{0} + \binom{m-1}{1} \right] (N+(i+1)\beta)^{[m-1]} (Np+i\beta) \\ & \quad + \left[\binom{m-1}{1} + \binom{m-1}{2} \right] (N+(i+2)\beta)^{[m-2]} (Np+i\beta)^{[2]} + \dots \end{aligned}$$

$$\begin{aligned}
& + (-1)^{m-2} \left[\binom{m-1}{m-3} + \binom{m-1}{m-2} \right] (N+(i+m-2)\beta)^{[2]} (Np+i\beta)^{[m-2]} \\
& + (-1)^{m-1} \left[\binom{m-1}{m-2} + \binom{m-1}{m-1} \right] (N+(i+m-1)\beta) (Np+i\beta)^{[m-1]} \\
& + (-1)^m (Np+i\beta)^{[m]} \\
& = (N+(i+1)\beta)^{[m-1]} (N+i\beta-Np-i\beta) - \binom{m-1}{1} (N+(i+2)\beta)^{[m-2]} (Np+i\beta) \\
& \quad \cdot [N+(i+1)\beta-Np-(i+1)\beta] + \dots \\
& + (-1)^{m-2} \binom{m-1}{m-2} (N+(i+m-1)\beta) (Np+i\beta)^{[m-2]} [N+(i+m-2)\beta-Np \\
& \quad - (i+m-2)\beta] \\
& + (-1)^{m-1} (Np+i\beta)^{[m-1]} (N+(i+m-1)\beta-Np-(i+m-1)\beta) \\
& = Nq [(N+(i+1)\beta)^{[m-1]} - \binom{m-1}{1} (N+(i+2)\beta)^{[m-2]} (Np+i\beta) + \dots \\
& \quad + (-1)^{m-2} \binom{m-1}{m-2} (N+(i+m-1)\beta) (Np+i\beta)^{[m-2]} \\
& \quad + (-1)^{m-1} (Np+i\beta)^{[m-1]}],
\end{aligned}$$

which by (2.4.9), with $k = m-1$ and $r = 1$, equals $Nq (Nq + \beta)^{[m-1]}$

$$= (Nq)^{[m]}.$$

Assume it is true for $k = m$, $r = s$.

Let $k = m$, $r = s+1$, then

$$\begin{aligned}
& \sum_{\ell=0}^m (-1)^\ell \binom{m}{\ell} (N+(i+s+1+\ell)\beta)^{[m-\ell]} (Np+i\beta)^{[\ell]} \\
&= (N+(i+s+1)\beta)^{[m]} - \binom{m}{1} (N+(i+s+2)\beta)^{[m-1]} (Np+i\beta) \\
&+ \binom{m}{2} (N+(i+s+3)\beta)^{[m-2]} (Np+i\beta)^{[2]} + \dots \\
&+ (-1)^{m-2} \binom{m}{m-2} (N+(i+s+m-1)\beta)^{[2]} (Np+i\beta)^{[m-2]} \\
&+ (-1)^{m-1} \binom{m}{m-1} (N+(i+s+m)\beta) (Np+i\beta)^{[m-1]} \\
&+ (-1)^m (Np+i\beta)^{[m]} \\
&= (N+(i+s+1)\beta)^{[m]} - \left[\binom{m-1}{0} + \binom{m-1}{1} \right] (N+(i+s+2)\beta)^{[m-1]} (Np+i\beta) \\
&+ \left[\binom{m-1}{1} + \binom{m-1}{2} \right] (N+(i+s+3)\beta)^{[m-2]} (Np+i\beta)^{[2]} + \dots \\
&+ (-1)^{m-2} \left[\binom{m-1}{m-3} + \binom{m-1}{m-2} \right] (N+(i+s+m-1)\beta)^{[2]} (Np+i\beta)^{[m-2]} \\
&+ (-1)^{m-1} \left[\binom{m-1}{m-2} + \binom{m-1}{m-1} \right] (N+(i+s+m)\beta) (Np+i\beta)^{[m-1]}
\end{aligned}$$

$$\begin{aligned}
& + (-1)^m (Np+i\beta)^{[m]} \\
& = (N+(i+s+2)\beta)^{[m-1]} (N+(i+s+1)\beta - Np - i\beta) \\
& \quad - \binom{m}{1} (N+(i+s+3)\beta)^{[m-3]} (Np+i\beta) (N+(i+s+2)\beta - Np - (i+1)\beta) + \dots \\
& \quad + (-1)^{m-2} \binom{m-1}{m-2} (N+(i+s+m)\beta) (Np+i\beta)^{[m-2]} (N+(i+s+m-1)\beta - Np - (i+m-2)\beta) \\
& \quad + (-1)^{m-1} (Np+i\beta)^{[m-1]} (N+(i+s+m)\beta - Np - (i+m-1)\beta) \\
& = (Nq+(s+1)\beta) [(N+(i+s+2)\beta)^{[m-1]} - \binom{m}{1} (N+(i+s+3)\beta)^{[m-3]} (Np+i\beta) \\
& \quad + \dots + (-1)^{m-2} \binom{m-1}{m-2} (N+(i+s+m)\beta) (Np+i\beta)^{[m-2]} \\
& \quad + (-1)^{m-1} (Np+i\beta)^{[m-1]}],
\end{aligned}$$

and by (2.4.9), with $k = m-1$ and $r = s+2$, equals

$$\begin{aligned}
& = (Nq+(s+1)\beta) (Nq+(s+2)\beta)^{[m-1]} \\
& = (Nq+(s+1)\beta)^{[m]}.
\end{aligned}$$

qed

Theorem 2.4.2 Consider two doubly indexed sequences $\{A_{ij}\}$ and $\{C_{ij}\}$ and positive integers W , B , and β , for which

$$\sum_{(i,j)} A_{ij} \frac{W[i] B[j]}{N[i+j]}$$

and

$$\sum_{(i,j)} C_{ij} \frac{W[i] B[j]}{N[i+j]}$$

are absolutely summable, and

$$|D_{ij}| = |A_{ij} - C_{ij}| \leq \frac{(1-\epsilon)^i (1-\eta)^j}{\frac{W[i]}{N[i]} \sum_{\ell=0}^j \binom{j}{\ell} \frac{(W+i\beta)[\ell]}{(N+i\beta)[\ell]}}$$

for $\epsilon, \eta > 0$. If

$$\sum_{(i,j)} A_{ij} p^i q^j = \sum_{(i,j)} C_{ij} p^i q^j \quad (2.4.10)$$

for $p \in (0, b)$, then

$$\sum_{(i,j)} A_{ij} \frac{W[i] B[j]}{N[i+j]} = \sum_{(i,j)} C_{ij} \frac{W[i] B[j]}{N[i+j]} .$$

Proof: As in the proof of Theorem 2.4.1, the absolute summability

of $\sum_{(i,j)} A_{ij} \frac{W[i] B[j]}{N[i+j]}$ and $\sum_{(i,j)} C_{ij} \frac{W[i] B[j]}{N[i+j]}$, (2.2.5), and

Fubini's Theorem (Theorem 19, page 269 of Royden [1968]) guarantee the absolute summability of $\sum_{(i,j)} A_{ij} p^i q^j$ and $\sum_{(i,j)} C_{ij} p^i q^j$ for almost all $p \in (0,1)$. Hence (2.4.10) implies

$$\sum_{i=0}^{\infty} \sum_{j=0}^{\infty} D_{ij} p^i q^j = 0 \quad (2.4.11)$$

for almost all $p \in (0,b)$. Substituting the binomial expansion of q^j in (2.4.11), for almost all $p \in (0,b)$,

$$\sum_{i=0}^{\infty} \sum_{j=0}^{\infty} D_{ij} p^i \sum_{k=0}^j \binom{j}{k} (-p)^k = 0. \quad (2.4.12)$$

In order to rearrange (2.4.12), it is necessary to establish the absolute summability of the left hand side:

$$\begin{aligned} & \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} |D_{ij}| \frac{W}{N} \frac{[i]}{[i]} \sum_{k=0}^j \binom{j}{k} \frac{(W+i\beta)^{[k]}}{(N+i\beta)^{[k]}} \\ & \leq \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \frac{1}{\frac{W}{N} \frac{[i]}{[i]} \sum_{\ell=0}^j \binom{j}{\ell} \frac{(W+i\beta)^{[\ell]}}{(N+i\beta)^{[\ell]}}} \frac{W}{N} \frac{[i]}{[i]} \sum_{\ell=0}^j \binom{j}{\ell} \frac{(W+i\beta)^{[\ell]}}{(N+i\beta)^{[\ell]}} \frac{(i-\epsilon)^i}{(1-\eta)^j} \\ & \leq \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} (1-\epsilon)^i (1-\eta)^j \\ & = \frac{1}{\epsilon\eta} \end{aligned} \quad (2.4.13)$$

For $p \leq \frac{W}{N}$, $p^i \leq \frac{W^{[i]}}{N^{[i]}}$, so (2.4.13) implies the absolute summability of (2.4.12) for $p \in (0, \frac{W}{N})$. Let $b' = \min(\frac{W}{N}, b)$, then for almost all $p \in (0, b')$,

$$\sum_{k=0}^{\infty} \left(\sum_{\ell=0}^k (-1)^{k-\ell} \sum_{j=k-\ell}^{\infty} \binom{j}{k-\ell} D_{\ell j} \right) p^k = 0. \quad (2.4.14)$$

Since the left hand side of (2.4.14) is a convergent power series for almost all $p \in (0, b')$, it is analytic and identically zero on $(0, b')$, which implies that it has derivatives of all orders on $[0, b']$, and that these are all zero on $[0, b']$, particularly at zero. Successive term-wise differentiation at zero (see page 198 of Buck [1965]) then yields

$$\sum_{\ell=0}^k (-1)^{k-\ell} \sum_{j=k-\ell}^{\infty} \binom{j}{k-\ell} D_{\ell j} = 0$$

for all $k = 0, 1, \dots$, and

$$\sum_{k=0}^{\infty} \left(\sum_{\ell=0}^k (-1)^{k-\ell} \sum_{j=k-\ell}^{\infty} \binom{j}{k-\ell} D_{\ell j} \right) \frac{W^{[k]}}{N^{[k]}} = 0. \quad (2.4.15)$$

The left hand side of (2.4.15) is a rearrangement of

$$\sum_{(i,j)} A_{ij} \frac{W^{[i]} B^{[j]}}{N^{[i+j]}} - \sum_{(i,j)} C_{ij} \frac{W^{[i]} B^{[j]}}{N^{[i+j]}} ,$$

which can be seen as follows:

Rewriting the left hand side of (2.4.15) as

$$\sum_{i=0}^{\infty} \sum_{j=0}^{\infty} D_{ij} \frac{W^{[i]}}{N^{[i]}} \sum_{\ell=0}^j \binom{j}{\ell} (-1)^{\ell} \frac{(W+i\theta)^{[\ell]}}{(N+i\beta)^{[\ell]}}, \quad (2.4.16)$$

we recognize that (2.4.16) equals zero by the absolute summability shown in (2.4.13). Multiplying

$$\sum_{\ell=0}^j \binom{j}{\ell} (-1)^{\ell} \frac{(W+i\beta)^{[\ell]}}{(N+i\beta)^{[\ell]}}$$

by $(N+i\beta)^{[j]}$ in (2.4.16) and applying Lemma 2.4.2, we have

$$\sum_{i=0}^{\infty} \sum_{j=0}^{\infty} D_{ij} \frac{W^{[i]} B^{[j]}}{N^{[i+j]}} = 0;$$

the absolute summability of $\sum_{(i,j)} A_{ij} \frac{W^{[i]} B^{[j]}}{N^{[i+j]}}$ and $\sum_{(i,j)} C_{ij} \frac{W^{[i]} B^{[j]}}{N^{[i+j]}}$

implies that

$$\sum_{(i,j)} A_{ij} \frac{W^{[i]} B^{[j]}}{N^{[i+j]}} = \sum_{(i,j)} C_{ij} \frac{W^{[i]} B^{[j]}}{N^{[i+j]}}.$$

qed

In Theorems 2.4.1 and 2.4.2, (2.4.1) holding for an interval of p ($(0,1)$ in Theorem 2.4.1 and $(0,b)$, $0 < b < 1$ in Theorem 2.4.2) implies (2.4.2) for all possible values of W , B , and β . Now, in studying the reverse implication, in Theorems 2.4.3 and 2.4.4, the more restrictive hypothesis for (2.4.2) holding on the tail of subsequence $\{(W_u, B_u)\}$ for which $\frac{W_u}{N_u}$ converges to a particular p^* , leads to a more restrictive conclusion: that (2.4.1) holds for the p^* only.

The two theorems differ in that, while Theorem 2.4.3 (similarly, Theorems 2.4.1 and 2.4.2) is stated in terms of an arbitrary order of summation, and hence, the necessity for the assumption of absolute summability of each side in (2.4.3), Theorem 2.4.4 is concerned only with a particular order of summation.

Theorem 2.4.3 Consider $p^* \in (0,1)$ as well as a positive integer β , and a sequence $\{W_u, B_u\}$ of pairs of positive integers such that $W_u \rightarrow +\infty$, $\frac{W_u}{N_u} \rightarrow p^*$, where $N_u = W_u + B_u$. Suppose that there are two doubly indexed sequences $\{A_{ij}\}$ and $\{C_{ij}\}$, and an integer m such that

$$(i) \quad \sum_{(i,j)} A_{ij} \frac{W_u^{[i]} B_u^{[j]}}{N_u^{[i+j]}}$$

is absolutely summable for $u \geq m$,

$$(ii) \quad \sum_{(i,j)} C_{ij} \frac{w_u^{[i]} B_u^{[j]}}{N_u^{[i+j]}}$$

is absolutely summable for $u \geq m$,

(iii) for $r = 1 + p$,

$\sum_{(i,j)} A_{ij} p^{irj}$ is uniformly absolutely summable for

$p \in (0,1)$: that is, there exists an $M < \infty$ such that

$$\sum_{(i,j)} |A_{ij}| p^{irj} \leq M$$

for $p \in (0,1)$, and

$$(iv) \quad \sum_{(i,j)} C_{ij} p^{irj}$$

is uniformly absolutely summable for $p \in (0,1)$.

Then if there exist an m' such that

$$\sum_{(i,j)} A_{ij} \frac{w_u^{[i]} B_u^{[j]}}{N_u^{[i+j]}} = \sum_{(i,j)} C_{ij} \frac{w_u^{[i]} B_u^{[j]}}{N_u^{[i+j]}} \quad (2.4.17)$$

for $u \geq m'$, then

$$\sum_{(i,j)} A_{ij} p^{*i} q^{*j} = \sum_{(i,j)} C_{ij} p^{*i} q^{*j}. \quad (2.4.18)$$

Proof: To begin with, in view of (i), (ii) and (2.2.5), we have for every $u \geq m$ that

$$\sum_{(i,j)} \int_0^1 |A_{ij}| p^i q^j \mathcal{B}_{\frac{w_u}{\beta}, \frac{B_u}{\beta}}(p) dp < +\infty,$$

and

$$\sum_{(i,j)} \int_0^1 |C_{ij}| p^i q^j \mathcal{B}_{\frac{w_u}{\beta}, \frac{B_u}{\beta}}(p) dp < +\infty.$$

By Fubini's Theorem (see Theorem 19, page 269 of Royden [1968]), (2.2.5), and (2.4.17) we have for every $u \geq \max(m, m')$,

$$\begin{aligned} & \int_0^1 \left[\sum_{(i,j)} A_{ij} p^i q^j \right] \mathcal{B}_{\frac{w_u}{\beta}, \frac{B_u}{\beta}}(p) dp \\ &= \sum_{(i,j)} A_{ij} \int_0^1 p^i q^j \mathcal{B}_{\frac{w_u}{\beta}, \frac{B_u}{\beta}}(p) dp \\ &= \sum_{(i,j)} A_{ij} \frac{w_u^{[i]} B_u^{[j]}}{N_u^{[i+j]}} \\ &= \sum_{(i,j)} C_{ij} \frac{w_u^{[i]} B_u^{[j]}}{N_u^{[i+j]}} \end{aligned}$$

$$\begin{aligned}
&= \sum_{(i,j)} c_{ij} \int_0^1 p^i q^j \mathcal{B}_{\frac{w_u}{\beta}, \frac{B_u}{\beta}}(p) dp \\
&= \int_0^1 \left[\sum_{(i,j)} c_{ij} p^i q^j \right] \mathcal{B}_{\frac{w_u}{\beta}, \frac{B_u}{\beta}}(p) dp,
\end{aligned}$$

so that, setting $f(p) = \sum_{(i,j)} A_{ij} p^i q^j - \sum_{(i,j)} c_{ij} p^i q^j$,

$$\int_0^1 f(p) \mathcal{B}_{\frac{w_u}{\beta}, \frac{B_u}{\beta}}(p) dp = 0. \quad (2.4.19)$$

In view of (iii) $\sum_{(i,j)} A_{ij} p^i (1-p)^j$ is absolutely summable for every $p \in (0,1)$, and thus can be rearranged into an absolutely convergent power series on $(0,1)$:

$$\sum_k a_k p^k = \sum_{(i,j)} A_{ij} p^i q^j = f_A(p).$$

But, in view of Theorem 22, p. 196 of Buck (1965), a power series is analytic (and hence continuous) within its circle of convergence, so that $f_A(p)$ is continuous on $(0,1)$.

The analogous remarks apply for $f_C(p) \equiv \sum_{(i,j)} c_{ij} p^i q^j$, so that the function $f(p) = f_A(p) - f_C(p)$ is continuous on $(0,1)$. Hence, given $\epsilon > 0$, there is a neighborhood m^* of p^* such that

$$|f(p) - f(p^*)| < \epsilon \quad (2.4.20)$$

for $p \in m^*$; also, in view of (iii) and (iv), there is a constant M^* such that

$$|f(p)| < M^* \quad (2.4.21)$$

for $p \in (0,1)$.

Now choose u greater than m such that

for $u \geq m$,

$$\int_{m^*} \mathcal{B}_{\frac{u}{\beta}, \frac{B_u}{\beta}}(p) dp \geq 1 - \epsilon, \quad (2.4.22)$$

and write (2.4.19) in the form

$$\begin{aligned} & \int_{m^*} f(p^*) \mathcal{B}_{\frac{u}{\beta}, \frac{B_u}{\beta}}(p) dp + \int_{m^*} (f(p) - f(p^*)) \mathcal{B}_{\frac{u}{\beta}, \frac{B_u}{\beta}}(p) dp \\ & + \int_{(0,1)-m^*} f(p) \mathcal{B}_{\frac{u}{\beta}, \frac{B_u}{\beta}}(p) dp = 0, \end{aligned}$$

and, taking the second and third terms to the right hand side,

$$\int_{m^*} f(p^*) \mathcal{B}_{\frac{w_u}{\beta}, \frac{B_u}{\beta}}(p) dp = - \int_{m^*} (f(p) - f(p^*)) \mathcal{B}_{\frac{w_u}{\beta}, \frac{B_u}{\beta}}(p) dp$$

$$- \int_{(0,1)-m^*} f(p) \mathcal{B}_{\frac{w_u}{\beta}, \frac{B_u}{\beta}}(p) dp;$$

then

$$f(p^*) = - \frac{\int_{m^*} (f(p) - f(p^*)) \mathcal{B}_{\frac{w_u}{\beta}, \frac{B_u}{\beta}}(p) dp - \int_{(0,1)-m^*} f(p) \mathcal{B}_{\frac{w_u}{\beta}, \frac{B_u}{\beta}}(p) dp}{\int_{m^*} \mathcal{B}_{\frac{w_u}{\beta}, \frac{B_u}{\beta}}(p) dp},$$

so that

$$|f(p^*)| \leq \frac{\int_{m^*} |f(p) - f(p^*)| \mathcal{B}_{\frac{w_u}{\beta}, \frac{B_u}{\beta}}(p) dp + \int_{(0,1)-m^*} |f(p)| \mathcal{B}_{\frac{w_u}{\beta}, \frac{B_u}{\beta}}(p) dp}{\int_{m^*} \mathcal{B}_{\frac{w_u}{\beta}, \frac{B_u}{\beta}}(p) dp}$$

$$\leq \frac{\epsilon + M^* \epsilon}{1 - \epsilon},$$

which follows from (2.4.20), (2.4.21), and (2.4.22).

qed

The situation when the A_{ij} 's and C_{ij} 's are bounded is given in Corollary 2.4.3a, where conditions (iii) and (iv) must be assumed. It is to be noted that if one assumes that the coefficients of terms containing p only or q only are zero, conditions (iii) and (iv) are satisfied, and Corollary 2.4.3a, itself, has a corollary, given in Corollary 2.4.3b.

Corollary 2.4.3a Consider a positive integer β and a sequence $\{W_u, B_u\}$ of pairs of positive integers such that $W_u \rightarrow \infty$, and $\frac{W_u}{N_u} \rightarrow p^* \in (0,1)$. Suppose that there are two doubly indexed sequences $\{A_{ij}\}$ and $\{C_{ij}\}$ for which $\sup_{(i,j)} |A_{ij}|, \sup_{(i,j)} |C_{ij}| \leq K < \infty$, and integer m such that

$$\sum_{(i,j)} A_{ij} \frac{W_u^{[i]} B_u^{[j]}}{N_u^{[i+j]}} = \sum_{(i,j)} C_{ij} \frac{W_u^{[i]} B_u^{[j]}}{N_u^{[i+j]}}$$

for $u \geq m$. Suppose also that conditions (iii) and (iv) of Theorem 2.4.3 hold, then

$$\sum_{(i,j)} A_{ij} p^{*i} q^{*j} = \sum_{(i,j)} C_{ij} p^{*i} q^{*j}.$$

Proof: Let $\sup_{(i,j)} |A_{ij}|, \sup_{(i,j)} |C_{ij}| \leq K < \infty$. Then $\sum_{(i,j)} |A_{ij}| p^i q^j \leq K \frac{1}{pq}$, and since we can choose m such that $W_u, B_u > \beta$ for $u \geq m$,

$$\begin{aligned}
& \int_0^1 \sum_{(i,j)} |A_{ij}| p^{iqj} \mathcal{B}_{\frac{W_u}{\beta}, \frac{B_u}{\beta}}(p) dp \\
& \leq K \int_0^1 \frac{1}{pq} \mathcal{B}_{\frac{W_u}{\beta}, \frac{B_u}{\beta}}(p) dp \\
& = K \frac{(N_u - \beta)(N_u - 2\beta)}{(W_u - \beta)(B_u - \beta)} < \infty.
\end{aligned}$$

This implies

$$\begin{aligned}
\sum_{(i,j)} |A_{ij}| \frac{W_u^{[i]} B_u^{[j]}}{N_u^{[i+j]}} &= \sum_{(i,j)} \int_0^1 |A_{ij}| p^{iqj} \mathcal{B}_{\frac{W_u}{\beta}, \frac{B_u}{\beta}}(p) dp \\
&\leq \infty,
\end{aligned}$$

so condition (i) of Theorem 2.4.3 holds, and a similar argument establishes condition (ii)

qed

Corollary 2.4.3b Consider a positive integer β and a sequence

$\{W_u, B_u\}$ of pairs of positive integers such that $W_u \rightarrow +\infty$ and

$\frac{W_u}{N_u} \rightarrow p^* \in (0,1)$. Suppose there are two doubly indexed sequences

$\{A_{ij}\}$ and $\{C_{ij}\}$ for which $\sup_{(i,j)} |A_{ij}|, \sup_{(i,j)} |C_{ij}| \leq K < \infty$, $A_{ij} = 0$

and $C_{ij} = 0$ if either $i = 0$ or $j = 0$, and

$$\sum_{(i,j)} A_{ij} \frac{w_u^{[i]} B_u^{[j]}}{N_u^{[i+j]}} = \sum_{(i,j)} C_{ij} \frac{w_u^{[i]} B_u^{[j]}}{N_u^{[i+j]}}$$

for all u . Then

$$\sum_{(i,j)} A_{ij} p^{*i} q^{*j} = \sum_{(i,j)} C_{ij} p^{*i} q^{*j}$$

Proof: Note that

$$\begin{aligned} & \int_0^1 \sum_{(i,j)} p^i q^j \mathcal{B}_{\frac{w_u}{\beta}, \frac{B_u}{\beta}}(p) dp \\ &= \int_0^1 \left(\sum_{i=1}^{\infty} p^i \right) \left(\sum_{j=1}^{\infty} q^j \right) \mathcal{B}_{\frac{w_u}{\beta}, \frac{B_u}{\beta}}(p) dp \\ &= \int_0^1 \left(\frac{p}{q} \right) \left(\frac{q}{p} \right) \mathcal{B}_{\frac{w_u}{\beta}, \frac{B_u}{\beta}}(p) dp \\ &= 1. \end{aligned}$$

Let $\sup_{(i,j)} |A_{ij}|, \sup_{(i,j)} |C_{ij}| \leq K$, then

$$K \geq \int_0^1 \sum_{(i,j)} |A_{ij}| p^i q^j \mathcal{B}_{\frac{w_u}{\beta}, \frac{B_u}{\beta}}(p) dp$$

$$= \sum_{(i,j)} \int_0^1 |A_{ij}| p^i q^j \mathcal{B}_{\frac{W_u}{\beta}, \frac{B_u}{\beta}}(p) dp,$$

so that condition (i) of Theorem 2.4.3 is satisfied with $m = 1$, and similarly for condition (ii).

Conditions (iii) and (iv) are checked in the same manner:
for example,

$$\begin{aligned} \sum_{(i,j)} |A_{ij}| p^i q^j &\leq K \left(\sum_{i=1}^{\infty} p^i \right) \left(\sum_{j=1}^{\infty} q^j \right) \\ &= K \left(\frac{p}{q} \right) \left(\frac{q}{p} \right) = K. \end{aligned}$$

qed

Theorem 2.4.4 Consider $p^* \in (0,1)$, as well as a positive integer β and a sequence $\{W_u, B_u\}$ of pairs of positive integers such that $W_u \rightarrow +\infty$, and $\frac{W_u}{N_u} \rightarrow p^*$. Suppose that there are two sequences $\{A_v\}$ and $\{C_v\}$, and an ordering $\{(i_v, j_v)\}$ of the points of $I \times I$, where I is the set of positive integers, such that

(i) There is an integer m such that

$$\sum_v A_v \frac{W_u^{[i_v]} B_u^{[j_v]}}{N_u^{[i_v+j_v]}}$$

converges uniformly in u , for $u \geq m$; that is given $\epsilon > 0$,

there is an $v_1(\epsilon)$ such that

$$\left| \sum_{v=k}^{\infty} A_v \frac{W_u^{[i_v]} B_u^{[j_v]}}{N_u^{[i_v+j_v]}} \right| \leq \epsilon \quad (2.4.23)$$

for all $u \geq m$ and $k \geq v_1(\epsilon)$;

(ii) There is an integer m such that

$$\sum_v C_v \frac{W_u^{[i_v]} B_u^{[j_v]}}{N_u^{[i_v+j_v]}}$$

converges uniformly in u , for $u \geq m$;

(iii) $\sum_v A_v p_v^{*i_v} q_v^{*j_v}$ converges;

(iv) $\sum_v C_v p_v^{*i_v} q_v^{*j_v}$ converges.

If there exists an m' such that

$$\sum_v A_v \frac{W_u^{[i_v]} B_u^{[j_v]}}{N_u^{[i_v+j_v]}} = \sum_v C_v \frac{W_u^{[i_v]} B_u^{[j_v]}}{N_u^{[i_v+j_v]}} \quad (2.4.24)$$

for $u \geq m'$, then

$$\sum_v A_v p_v^{*i_v} q_v^{*j_v} = \sum_v C_v p_v^{*i_v} q_v^{*j_v}.$$

Proof: By assumptions (iii) and (iv), given $\epsilon > 0$, there is a $v_2(\epsilon)$ such that, for $k \geq v_2(\epsilon)$,

$$\left| \sum_{v=k}^{\infty} A_v p_v^{* i_v} q_v^{* j_v} \right| \leq \epsilon \quad (2.4.25)$$

and

$$\left| \sum_{v=k}^{\infty} C_v p_v^{* i_v} q_v^{* j_v} \right| < \epsilon. \quad (2.4.26)$$

Finally, it is clear, in view of (2.4.24), given $\epsilon > 0$, and letting $v_0(\epsilon) = \max \{v_1(\epsilon), v_2(\epsilon)\}$, that there exists an $m(\epsilon)$ such that

$$\left| \sum_{v=0}^{v_0(\epsilon)} A_v \frac{w_u^{[i_v]} B_u^{[j_v]}}{N_u^{[i_v+j_v]}} - \sum_{v=0}^{v_0(\epsilon)} A_v p_v^{* i_v} q_v^{* j_v} \right| \leq \epsilon \quad (2.4.27)$$

for $u \geq m(\epsilon)$, and similarly for the C_v 's.

Hence, given $\epsilon > 0$, write, for any $u > \max \{m, m', m(\epsilon)\}$,

$$\begin{aligned} & \left| \sum_v A_v p_v^{* i_v} q_v^{* j_v} - \sum_v C_v p_v^{* i_v} q_v^{* j_v} \right| \\ &= \left| \sum_{v=1}^{v_0(\epsilon)} A_v p_v^{* i_v} q_v^{* j_v} + \sum_{v=v_0(\epsilon)+1}^{\infty} A_v p_v^{* i_v} q_v^{* j_v} \right. \\ & \quad \left. - \sum_{v=1}^{v_0(\epsilon)} C_v p_v^{* i_v} q_v^{* j_v} - \sum_{v=v_0(\epsilon)+1}^{\infty} C_v p_v^{* i_v} q_v^{* j_v} \right| \end{aligned}$$

$$\begin{aligned}
&= \left| \begin{aligned} &v_0(\epsilon) \quad i_v \quad j_v \quad v_0(\epsilon) \\ &\sum_{v=1}^{\infty} A_v p^* \quad q^* \quad - \sum_{v=1}^{\infty} A_v \frac{w_u[i_v] B_u[j_v]}{N_u[i_v+j_v]} \end{aligned} \right| \\
&\quad - \sum_{v=1}^{\infty} \frac{v_0(\epsilon)}{C_v p^*} i_v \quad j_v + \sum_{v=1}^{\infty} \frac{v_0(\epsilon)}{C_v} \frac{w_u[i_v] B_u[j_v]}{N_u[i_v+j_v]} \\
&\quad + \sum_{v=v_0(\epsilon)+1}^{\infty} A_v p^* i_v \quad q^* j_v - \sum_{v=v_0(\epsilon)+1}^{\infty} C_v p^* i_v \quad q^* j_v \\
&\quad - \sum_{v=v_0(\epsilon)+1}^{\infty} A_v \frac{w_u[i_v] B_u[j_v]}{N_u[i_v+j_v]} + \sum_{v=v_0(\epsilon)+1}^{\infty} C_v \frac{w_u[i_v] B_u[j_v]}{N_u[i_v+j_v]} \\
&\quad + \sum_v A_v \frac{w_u[i_v] B_u[j_v]}{N_u[i_v+j_v]} - \sum_v C_v \frac{w_u[i_v] B_u[j_v]}{N_u[i_v+j_v]} \\
&\leq \left| \begin{aligned} &v_0(\epsilon) \quad i_v \quad j_v \quad v_0(\epsilon) \\ &\sum_{v=1}^{\infty} A_v p^* \quad q^* - \sum_{v=1}^{\infty} A_v \frac{w_u[i_v] B_u[j_v]}{N_u[i_v+j_v]} \end{aligned} \right| \\
&\quad + \left| \begin{aligned} &v_0(\epsilon) \quad i_v \quad j_v \quad v_0(\epsilon) \\ &\sum_{v=1}^{\infty} C_v p^* \quad q^* - \sum_{v=1}^{\infty} C_v \frac{w_u[i_v] B_u[j_v]}{N_u[i_v+j_v]} \end{aligned} \right| \\
&\quad + \left| \sum_{v=v_0(\epsilon)+1}^{\infty} A_v p^* i_v \quad q^* j_v \right|
\end{aligned}$$

$$\begin{aligned}
& + \left| \sum_{v=v_0(\epsilon)+1}^{\infty} C_v p_v^{* i_v} q_v^{* j_v} \right| \\
& + \left| \sum_{v=v_0(\epsilon)+1}^{\infty} A_v \frac{W_u^{[i_v]} B_u^{[j_v]}}{N_u^{[i_v+j_v]}} \right| \\
& + \left| \sum_{v=v_0(\epsilon)+1}^{\infty} C_v \frac{W_u^{[i_v]} B_u^{[j_v]}}{N_u^{[i_v+j_v]}} \right| \\
& + \left| \sum_v A_v \frac{W_u^{[i_v]} B_u^{[j_v]}}{N_u^{[i_v+j_v]}} - \sum_v C_v \frac{W_u^{[i_v]} B_u^{[j_v]}}{N_u^{[i_v+j_v]}} \right| \\
& \leq \epsilon + \epsilon + \epsilon + \epsilon + \epsilon + \epsilon + 0.
\end{aligned}$$

The six ϵ 's are due, respectively, to (2.4.27), the analog of (2.4.27) for the C_{ij} 's, (2.4.25), (2.4.26), (2.4.23), and the C_{ij} analog of (2.4.23), and the zero is due to (2.4.24).

The case when the coefficients are bounded is given by

Corollary 2.4.4a Consider a positive integer β and a sequence $\{W_u, B_u\}$ of pairs of positive integers such that $W_u \rightarrow +\infty$ and

$\frac{W_u}{N_u} \rightarrow +\infty$ and $\frac{W_u}{N_u} \rightarrow p^* \in (0,1)$. Suppose there are two sequences $\{A_v\}$

and $\{C_v\}$, for which $\sup_v |A_v|, \sup_v |C_v| \leq K < \infty$, an ordering $\{(i_v \times j_v)\}$ of the points $I \times I$, and an integer m such that

$$\sum_v \frac{W_u^{[i_v]} B_u^{[j_v]}}{N_u^{[i_v+j_v]}}$$

is uniformly convergent in u , for $u \geq m$, and

$$\sum_v A_v \frac{W_u^{[i_v]} B_u^{[j_v]}}{N_u^{[i_v+j_v]}} = \sum_v C_v \frac{W_u^{[i_v]} B_u^{[j_v]}}{N_u^{[i_v+j_v]}}$$

for $u \geq m$. Then

$$\sum_v A_v p_v^{*i_v} q_v^{*j_v} = \sum_v C_v p_v^{*i_v} q_v^{*j_v}.$$

Proof: The boundedness of the A_v 's and C_v 's insures the convergence of $\sum_v A_v p_v^{*i_v} q_v^{*j_v}$ and $\sum_v C_v p_v^{*i_v} q_v^{*j_v}$.

qed

If the summation is taken over a finite number of terms, all conditions of the previous four theorems are trivially satisfied.

Corollary 2.4.4b Let n be finite and A_{ij} , C_{ij} be constants.

Then

$$\sum_{i=0}^n \sum_{j=0}^n A_{ij} \frac{W^{[i]} B^{[j]}}{N^{[i+j]}} = \sum_{i=0}^n \sum_{j=0}^n C_{ij} \frac{W^{[i]} B^{[j]}}{N^{[i+j]}}$$

for all positive integers W , B , β , and n , if and only if

$$\sum_{i=0}^n \sum_{j=0}^n A_{ij} p^i q^j = \sum_{i=0}^n \sum_{j=0}^n C_{ij} p^i q^j$$

for all $p \in (0,1)$.

We now turn our attention to the multivariate version of the relation between identities of sums of "Polya-like" terms and the corresponding identities of "Bernoulli-like" terms. As in the second section of the chapter, the results will be given for a population consisting of k types.

We ask under what conditions

$$\sum_{(i_1, \dots, i_k)} A_{i_1 \dots i_k} \prod_{j=1}^k p_j^{i_j} = \sum_{(i_1, \dots, i_k)} C_{i_1 \dots i_k} \prod_{j=1}^k p_j^{i_j}$$

implies, and in turn is implied by

$$\sum_{(i_1, \dots, i_k)} A_{i_1 \dots i_k} \frac{\prod_{j=1}^k W_j^{[i_j]}}{N^{[\sum_{j=1}^k i_j]}} = \sum_{(i_1, \dots, i_k)} C_{i_1 \dots i_k} \frac{\prod_{j=1}^k W_j^{[i_j]}}{N^{[\sum_{j=1}^k i_j]}}$$

for arbitrary constants $A_{i_1 \dots i_k}$ and $C_{i_1 \dots i_k}$.

Theorems 2.4.1 and 2.4.4 are the most readily generalized, and are given here as Theorems 2.4.1a and 2.4.4a with their respective corollaries for bounded coefficients. Corollary 2.4.5a covers the case of finite sums.

Theorem 2.4.1a Assume that $\sum_{(i_1, \dots, i_k)} A_{i_1 \dots i_k} \prod_{j=1}^k p_j^{i_j}$ and $\sum_{(i_1, \dots, i_k)} C_{i_1 \dots i_k} \prod_{j=1}^k p_j^{i_j}$ are absolutely summable and finite and that

$$\sum_{(i_1, \dots, i_k)} A_{i_1 \dots i_k} \prod_{j=1}^k p_j^{i_j} = \sum_{(i_1, \dots, i_k)} C_{i_1 \dots i_k} \prod_{j=1}^k p_j^{i_j}$$

for all $p_1, \dots, p_k \in (0, 1)$ such that $\sum_{i=1}^k p_i = 1$. Then

$$\begin{aligned} & \sum_{(i_1, \dots, i_k)} A_{i_1 \dots i_k} \frac{\prod_{j=1}^k w_j^{[i_j]}}{\left[\sum_{j=1}^k i_j \right]_N} \\ &= \sum_{(i_1, \dots, i_k)} C_{i_1 \dots i_k} \frac{\prod_{j=1}^k w_j^{[i_j]}}{\left[\sum_{j=1}^k i_j \right]_N} \end{aligned}$$

for all positive integers w_1, \dots, w_k and β such that $N = \sum_{i=1}^k w_i$.

Corollary 2.4.1b If the $A_{i_1 \dots i_k}$'s and $C_{i_1 \dots i_k}$'s are bounded in absolute value, and

$$\sum_{(i_1, \dots, i_k)} A_{i_1 \dots i_k} \prod_{j=1}^k p_j^{i_j} = \sum_{(i_1, \dots, i_k)} C_{i_1 \dots i_k} \prod_{j=1}^k p_j^{i_j}$$

for all $p_1, \dots, p_k \in (0,1)$ such that $\sum_{i=1}^k p_i = 1$,

then

$$\sum_{(i_1, \dots, i_k)} A_{i_1 \dots i_k} \frac{\prod_{j=1}^k w_j^{[i_j]}}{\left[\sum_{j=1}^k i_j \right]_N} = \sum_{(i_1, \dots, i_k)} C_{i_1 \dots i_k} \frac{\prod_{j=1}^k w_j^{[i_j]}}{\left[\sum_{j=1}^k i_j \right]_N}$$

for all positive integers w_1, \dots, w_k , and β .

Theorem 2.4.4a Fix β and choose $p_1, \dots, p_k \in (0,1)$ such that

$\sum_{i=1}^k p_i = 1$. Let $R(m)$ be the set of all k -tuples (i_1, \dots, i_k) such

that $\max(i_1, \dots, i_k) = m$. Suppose there exist subsequences of positive integers $\{w_1, \dots, w_k\}$ such that $\frac{w_1}{N} \rightarrow p_1, \dots, \frac{w_k}{N} \rightarrow p_k$,

and for which

$$\sum_{m=0}^{\infty} \sum_{(i_1, \dots, i_k) \in R(m)} A_{i_1 \dots i_k} \frac{\prod_{j=1}^k w_j^{[i_j]}}{\left[\sum_{j=1}^k i_j \right]_N} =$$

$$\sum_{m=0}^{\infty} \sum_{(i_1, \dots, i_k) \in R(m)} C_{i_1 \dots i_k} \frac{\prod_{j=1}^k w_j^{[i_j]}}{N^{\frac{k}{\sum_{j=1}^k i_j}}}$$

exists and is finite. If, given $\varepsilon > 0$, there exist an m_0 such that

$$\sum_{m=m_0}^{\infty} \sum_{(i_1, \dots, i_k) \in R(m)} A_{i_1 \dots i_k} \frac{\prod_{j=1}^k w_j^{[i_j]}}{N^{\frac{k}{\sum_{j=1}^k i_j}}} \leq \frac{\varepsilon}{6}$$

and

$$\sum_{m=m_0}^{\infty} \sum_{(i_1, \dots, i_k) \in R(m)} C_{i_1 \dots i_k} \frac{\prod_{j=1}^k w_j^{[i_j]}}{N^{\frac{k}{\sum_{j=1}^k i_j}}} \leq \frac{\varepsilon}{6}$$

on $\{w_1, \dots, w_k\}$, and if $\sum_{(i_1, \dots, i_k)} A_{i_1 \dots i_k} \prod_{j=1}^k p_j^{i_j}$ and

$\sum_{(i_1, \dots, i_k)} C_{i_1 \dots i_k} \prod_{j=1}^k p_j^{i_j}$ are absolutely summable, then

$$\sum_{(i_1, \dots, i_k)} A_{i_1 \dots i_k} \prod_{j=1}^k p_j^{i_j} = \sum_{(i_1, \dots, i_k)} C_{i_1 \dots i_k} \prod_{j=1}^k p_j^{i_j}.$$

Corollary 2.4.4a Fix β and choose $p_1, \dots, p_k \in (0,1)$ such that

$\sum_{i=1}^k p_i = 1$. If the $A_{i_1 \dots i_k}$'s and $C_{i_1 \dots i_k}$'s are bounded in absolute value, and there exist subsequences of positive integers $\{w_1, \dots, w_k\}$ such that $\frac{w_1}{N} \rightarrow p_1, \dots, \frac{w_k}{N} \rightarrow p_k$, and for which

$$\sum_{m=0}^{\infty} \sum_{(i_1, \dots, i_k) \in R(m)} A_{i_1 \dots i_k} \frac{\prod_{j=1}^k w_j^{[i_j]}}{N^{\left[\sum_{j=1}^k i_j \right]}}$$

$$= \sum_{m=0}^{\infty} \sum_{(i_1, \dots, i_k) \in R(m)} C_{i_1 \dots i_k} \frac{\prod_{j=1}^k w_j^{[i_j]}}{N^{\left[\sum_{j=1}^k i_j \right]}} ,$$

then

$$\sum_{(i_1, \dots, i_k)} A_{i_1 \dots i_k} \prod_{j=1}^k p_j^{i_j} = \sum_{(i_1, \dots, i_k)} C_{i_1 \dots i_k} \prod_{j=1}^k p_j^{i_j}.$$

Corollary 2.4.4b Given constants $A_{i_1 \dots i_k}$ and $C_{i_1 \dots i_k}$ and finite n ,

$$\sum_{i_1, \dots, i_k=0}^n A_{i_1 \dots i_k} \prod_{j=1}^k p_j^{i_j} = \sum_{i_1, \dots, i_k=0}^n C_{i_1 \dots i_k} \prod_{j=1}^k p_j^{i_j},$$

for all $p_j \in (0,1)$ such that $\sum_{j=1}^k p_j = 1$, if and only if

$$\sum_{i_1, \dots, i_k=0}^n A_{i_1 \dots i_k} \frac{\prod_{j=1}^k w_j^{[i_j]}}{\left[\sum_{j=1}^k i_j \right]_N}$$

$$= \sum_{i_1, \dots, i_k=0}^n C_{i_1 \dots i_k} \frac{\prod_{j=1}^k w_j^{[i_j]}}{\left[\sum_{j=1}^k i_j \right]_N},$$

for all values of w_j , $j=1, \dots, k$.

III. CLOSURE

Recall from Chapter I that GMS stop rule partitions are characterized as boundaries in the (n,d) -plane. In the work of Girshick, Mosteller, and Savage (1946), the point (n,d) represents the occurrence of d white draws out of a total of n draws under Bernoulli-sampling. Since the points themselves have no probability structure, exchangeability allows the point (n,d) to be thought of, also, as representing a sample of d white and $n-d$ black balls resulting from either Polya or hypergeometric sampling.

We will also use the terms Bernoulli sampling and Polya sampling to describe the probability assigned to the points (n,d) generated by the Bernoulli and Polya processes, respectively.

Denote by B^* , the boundary corresponding to a GMS stop rule partition. Girshick, Mosteller, and Savage (1946) defined closure to mean $P_b(B^*) = 1$, and established a sufficient condition for its existence: $\lim_{n \rightarrow \infty} \frac{A(n)}{\sqrt{n}} = 0$, where $A(n)$ represents the number of points at each n for which sampling continues. Wolfowitz (1946) relaxed the condition to $\lim_{n \rightarrow \infty} \frac{A(n)}{\sqrt{n}} < \infty$.

In this chapter, we define closure to mean $P_u(S) = 1$, and establish a similar sufficient condition for closure for Polya sampling for general stop rules, which is

$$\lim_{n \rightarrow \infty} \frac{A(n)}{n} = 0,$$

where now $A(n)$ is the number of distinct values at time n of $\sum_{i=1}^n x_i$ for the path segments for which sampling continues.

It will also be shown that any stop rule closed under Bernoulli sampling is closed under Polya sampling, and that any stop rule not closed for an interval of p in the Bernoulli case, will not be closed for any values of the parameters W , B , and β in the Polya case.

For a GMS stop rule partition, we show that if the boundary can be approximated by diverging lines, then the stop rule is not closed under Polya sampling. As an example of this type of stop rule partition we derive the boundary of the SPRT under Polya sampling.

Consider, for a moment, a GMS stop rule partition. The set of points corresponding to the path segments for which sampling continues is called the sampling region R by Girshick, Mosteller, and Savage (1946). Suppose that the R lies between two curves; the two conditions for closure imply that for large n , the distance between the curves, the "throat," is of order $O(\sqrt{n})$ in the Bernoulli case and $o(n)$ in the Polya case. In Chapter II, it was noted that the variance of the Bernoulli process is a function of n , while that of the Polya process is a function of n^2 , so the larger Polya variance permits a larger sampling region be closed with respect to Polya sampling.

The first result of this chapter is the proof of the sufficient condition for closure under Polya sampling, which requires two lemmas

establishing the various shapes of the Polya density under different values of the parameters W , B , β , and now n .

Lemma 3.1 Let f_u be the Polya density with parameters n , W , B , $\beta > 0$. The shape of f_u is then described by one of the following five cases:

- (1) For large n , if $\frac{W}{\beta} \leq 1$ and $\frac{B}{\beta} \leq 1$, f_u has a discrete U-shaped density;
- (2) For large n , if $\frac{W}{\beta} > 1$ and $\frac{B}{\beta} > 1$, f_u is discrete and unimodal;
- (3) If $\frac{W}{\beta} = \frac{B}{\beta} = 1$, f_u has the discrete uniform density;
- (4) If $\frac{W}{\beta} \geq 1$ and $\frac{B}{\beta} \leq 1$, but $W \neq B$, then f_u is monotone increasing;
- (5) If $\frac{W}{\beta} \leq 1$ and $\frac{B}{\beta} \geq 1$, but $W \neq B$, then f_u is monotone decreasing.

Proof: Let f_u be a Polya density with n representing the number of white and black balls, respectively, in the urn at time zero, and β , the number of balls added after each trial.

$$\frac{P_u \left(\sum_{i=1}^n X_i = k+1 \right)}{P_u \left(\sum_{i=1}^n X_i = k \right)} = \frac{\binom{n}{k+1} W^{[k+1]} B^{[n-k-1]}}{\binom{n}{k} W^{[k]} B^{[n-k]}}$$

$$= \frac{n-k}{k+1} \cdot \frac{W+k\beta}{B+(n-k-1)\beta} = 1 + \frac{n(W-\beta) - (B-\beta) - k(N-2\beta)}{(k+1)(B+(n-k-1)\beta)},$$

$$k = 0, \dots, n-1.$$

Then f_u reaches a maximum or minimum at k , when the above ratio is equal to one; that is, when

$$k = \frac{n(W-\beta) - (B-\beta)}{(W-\beta) + (B-\beta)}.$$

Let $m(n) = \frac{n(W-\beta) - (B-\beta)}{(W-\beta) + (B-\beta)}$. If $m(n)$ exists, then f_u is increasing when

$$n(W-\beta) - (B-\beta) > k(N-2\beta),$$

and decreasing when

$$n(W-\beta) - (B-\beta) \leq k(N-2\beta).$$

The occurrence of these events depends on the values W , B , β assume.

Case 1. Let $\frac{W}{\beta} \leq 1$ and $\frac{B}{\beta} < 1$, then $W-\beta \leq 0$ and $B-\beta < 0$. There exist an n large enough such that $n|W-\beta| > |B-\beta|$, so

$$0 \leq m(n) = \frac{n(W-\beta) - (B-\beta)}{(W-\beta) + (B-\beta)} < n.$$

Then f_u increases when

$$n(W-\beta) - (B-\beta) > k(N-2\beta); \text{ or,}$$

since $(N-2\beta) \leq 0$, $k > m(n)$, and decreases when

$$n(W-\beta) - (B-\beta) \leq k(N-2\beta), \text{ or}$$

$$k \leq m(n).$$

Case 2. Let $\frac{W}{\beta} > 1$ and $\frac{B}{\beta} > 1$, then $W-\beta > 0$ and $B-\beta > 0$.

For n large enough, $n(W-\beta) > (B-\beta)$; so

$$0 \leq m(n) = \frac{n(W-\beta) - (B-\beta)}{(W-\beta) + (B-\beta)} \leq n.$$

Now f_u increases when $n(W-\beta) - (B-\beta) > k(N-2\beta)$, or $k \leq m(n)$,

and decreases when $n(W-\beta) - (B-\beta) \leq k(N-2\beta)$, or $k > m(n)$.

Case 3. Let $\frac{W}{\beta} = 1$ and $\frac{B}{\beta} = 1$, then

$$\begin{aligned} P_u \left(\sum_{i=1}^n X_i = k \right) &= \frac{\binom{n}{k} \prod_{i=0}^{k-1} \left(\frac{W}{\beta} + i \right) \prod_{i=0}^{n-k-1} \left(\frac{B}{\beta} + i \right)}{\prod_{i=0}^{n-1} \left(\frac{N}{\beta} + i \right)} \\ &= \frac{n!}{k! (n-k)!} \cdot \frac{k! (n-k)!}{(n+1)!} = \frac{1}{n+1} \end{aligned}$$

for all $k = 0, \dots, n$.

Case 4. Let $\frac{W}{\beta} \geq 1$, $\frac{B}{\beta} \leq 1$, and $W \neq B$.

(a) Suppose $\frac{W}{\beta} > 1$ and $\frac{B}{\beta} < 1$, then

$$\frac{P_u \left(\sum_{i=1}^n X_i = k+1 \right)}{P_u \left(\sum_{i=1}^n X_i = k \right)} = \frac{n-k}{k+1} \cdot \frac{\frac{W}{\beta} + k}{\frac{B}{\beta} + n-k-1}$$

$$> \frac{n-k}{k+1} \frac{k+1}{\frac{B}{p} + n-k-1} > \frac{n-k}{n-k} = 1.$$

(b) Suppose $\frac{W}{p} = 1$, $\frac{B}{p} \leq 1$, then

$$\frac{P_u \left(\sum_{i=1}^n X_i = k+1 \right)}{P_u \left(\sum_{i=1}^n X_i = k \right)} = \frac{n-k}{k+1} \frac{1+k}{\frac{B}{p} + n-k-1} > \frac{n-k}{1 + n-k-1} = 1.$$

(c) Suppose $\frac{W}{p} > 1$, $\frac{B}{p} = 1$ then

$$\frac{P_u \left(\sum_{i=1}^n X_i = k+1 \right)}{P_u \left(\sum_{i=1}^n X_i = k \right)} = \frac{n-k}{k+1} \frac{\frac{W}{p} + k}{1 + n-k-1} > \frac{1+k}{k+1} = 1.$$

Case 5. Let $\frac{W}{p} \leq 1$ and $\frac{B}{p} \geq 1$ with $W \neq B$.

(a) Suppose $\frac{W}{p} \leq 1$ and $\frac{B}{p} > 1$, then

$$\begin{aligned} \frac{P_u \left(\sum_{i=1}^n X_i = k+1 \right)}{P_u \left(\sum_{i=1}^n X_i = k \right)} &= \frac{n-k}{k+1} \cdot \frac{\frac{W}{p} + k}{\frac{B}{p} + n-k-1} < \frac{n-k}{k+1} \frac{1+k}{\frac{B}{p} + n-k} \\ &\leq \frac{n-k}{1 + n-k-1} = 1. \end{aligned}$$

(b) Suppose $\frac{W}{p} = 1$ and $\frac{B}{p} > 1$, then

$$\frac{P_u \left(\sum_{i=1}^n X_i = k+1 \right)}{P_u \left(\sum_{i=1}^n X_i = k \right)} = \frac{n-k}{k+1} \cdot \frac{1+k}{\frac{B}{\beta} + n-k-1} \leq \frac{n-k}{1+n-k-1} = 1.$$

(c) Suppose $\frac{W}{\beta} \leq 1$ and $\frac{B}{\beta} = 1$, then

$$\frac{P_u \left(\sum_{i=1}^n X_i = k+1 \right)}{P_u \left(\sum_{i=1}^n X_i = k \right)} = \frac{n-k}{k+1} \frac{\frac{W}{\beta} + k}{1+n-k-1} \leq \frac{1+k}{k+1} = 1.$$

Lemma 3.2 Let f_u be a Polya density with $\frac{W}{\beta} \leq 1$, $\frac{B}{\beta} \leq 1$ ($\frac{W}{\beta} > 1$, $\frac{B}{\beta} > 1$). Define $m(n)$ to be the minimum (maximum) of f_u . Then

$\lim_{n \rightarrow \infty} \frac{m(n)}{n} = m$, where m is the minimum (maximum) of $\mathcal{B}_{\frac{W}{\beta}, \frac{B}{\beta}}$, the

beta limiting distribution of $\frac{\sum_{i=1}^n X_i}{n}$.

Proof: By Lemma 3.1, $m(n)$ is the minimum (maximum) of f_u when $\frac{W}{\beta} \leq 1$ and $\frac{B}{\beta} < 1$ ($\frac{W}{\beta} > 1$ and $\frac{B}{\beta} > 1$), and

$$\frac{m(n)}{n} = \frac{W-\beta - \left(\frac{B-\beta}{n}\right)}{(W-\beta) + (B-\beta)} \xrightarrow{n \rightarrow \infty} \frac{W-\beta}{W+B-2\beta}.$$

The maximum or minimum of a beta density with parameters $\frac{W}{\beta}$ and $\frac{B}{\beta}$ is given by

$$\frac{d}{dz} \left[z^{\frac{W}{\beta}-1} (1-z)^{\frac{B}{\beta}-1} \right] = 0,$$

solved for z , so that

$$\left(\frac{W}{\beta} - 1\right) z^{\frac{W}{\beta}-2} (1-z)^{\frac{B}{\beta}-1} - \left(\frac{B}{\beta}-1\right) z^{\frac{W}{\beta}-1} (1-z)^{\frac{B}{\beta}-2} = 0,$$

and simplifying, we have

$$\left(\frac{W}{\beta}-1\right)(1-z) - \left(\frac{B}{\beta}-1\right) z = 0.$$

Thus,

$$m = \frac{\frac{W}{\beta}-1}{\left(\frac{W}{\beta}-1\right) + \left(\frac{B}{\beta}-1\right)},$$

so that

$$\frac{m(n)}{n} \xrightarrow{n \rightarrow \infty} m.$$

qed

We are now in a position to state and prove the sufficient condition for closure for a stop rule under Polya sampling.

Theorem 3.1 A sufficient condition that a stop rule S is closed under Polya sampling is that

$$\lim_{n \rightarrow \infty} \frac{A(n)}{n} = 0,$$

where $A(n)$ is the number of distinct values of $\sum_{i=1}^n X_i$ for the path segments for which sampling continues at time n .

Proof: Define χ to be the set of all finite sequences of 0's and 1's. At each n , the set of path segments for which sampling continues is given by the set

$$A_n = \{\underline{x}_n \in \chi: \underline{x}_n \not\in S, \text{ and for } 1 \leq k \leq n-1, \underline{x}_k \not\in S\}.$$

Define the set $b_n = \{\underline{x}_n \in S\}$ which is the set of sequences for which termination occurs at time n , then $\bigcup_{i=1}^{\infty} b_i = S$. Next, define $K_n = \bigcup_{i=1}^n b_i$, the set of path segments which have been stopped up to and including time n .

We want to show $P_u(S) = 1$. Since A_n is the set of all path segments at time n which have not been stopped previously, it is clear that

$$1 = P_u(K_n \cup A_n) = P_u(K_n) + P_u(A_n).$$

Now $\lim_{n \rightarrow \infty} P_u(K_n) = P_u(S)$, so that it remains to show that $\lim_{n \rightarrow \infty} P_u(A_n) = 0$.

Partition the set A_n into mutually exclusive sets A_n^i , where

$$A_n^i = \{\underline{x}_n \in A_n: \sum_{j=1}^n x_j = a_i\};$$

that is, A_n^i is the subset of A_n for which the first n draws yield

exactly a_i white balls. Let $A(n)$ be the number of distinct subsets A_n^i in A_n , then

$$A_n = \bigcup_{i=1}^{A(n)} A_n^i.$$

Next define the sets $A_n'^i$ as

$$A_n'^i = \{\underline{x}_n \in \chi: \sum_{j=1}^n x_j = a_i\}, i=1, \dots, A(n),$$

so $A_n^i \subset A_n'^i$. Now

$$P_u(A_n) = \sum_{i=1}^{A(n)} P_u(A_n^i) \leq \sum_{i=1}^{A(n)} P_u(A_n'^i) = \sum_{i=1}^{A(n)} P_u(\underline{x}_n \in \chi: \sum_{j=1}^n x_j = a_i).$$

Let W_n (B_n) be the number of white (black) balls in the urn after the n th draw. From Freedman (1965), we have that

$$\frac{W_n}{W_n + B_n} \xrightarrow[n \rightarrow \infty]{a.s.} Z_{\frac{W}{\beta}, \frac{B}{\beta}},$$

where $Z_{\frac{W}{\beta}, \frac{B}{\beta}}$ has the beta density, $\mathcal{B}_{\frac{W}{\beta}, \frac{B}{\beta}}$, given in (2.2.5), so that

for constants $0 \leq D_1 \leq D_2 \leq 1$,

$$P_u(D_1 \leq \frac{W_n}{W_n + B_n} \leq D_2) \xrightarrow[n \rightarrow \infty]{} \int_{D_1}^{D_2} \mathcal{B}_{\frac{W}{\beta}, \frac{B}{\beta}}(p) dp.$$

Since

$$\frac{W_n}{W_n + B_n} = \frac{W + \beta \sum_{i=1}^n X_i}{W + B + n\beta},$$

(3.1) becomes

$$P_u \left(\frac{D_1 (W + B + n\beta) - W}{\beta} \leq \sum_{i=1}^n X_i \leq \frac{D_2 (W + B + n\beta) - W}{\beta} \right) \\ \xrightarrow{n \rightarrow \infty} \int_{D_1}^{D_2} \mathcal{B}_{\frac{W}{\beta}, \frac{B}{\beta}}(p) dp.$$

The number of integers in the interval

$$\left[\frac{D_1 (W + B + n\beta) - W}{\beta}, \frac{D_2 (W + B + n\beta) - W}{\beta} \right]$$

is $\frac{(D_2 - D_1)(W + B)}{\beta} + n(D_2 - D_1)$, and hence of order n .

There are four parametric cases to consider.

Case 1: $\frac{W}{\beta} \ll 1$, $\frac{B}{\beta} \ll 1$. Choose a sequence of integers $\{K_1(n)\}$ and $\{K_2(n)\}$ such that $0 \leq K_1(n)$, $K_2(n) \leq n$, and $A(n) = K_1(n) + K_2(n)$ for all n . Since $\lim_{n \rightarrow \infty} \frac{A(n)}{n} = 0$, there exist a subsequence n_v such that $\frac{A(n_v)}{n_v} \xrightarrow{v \rightarrow \infty} 0$, so that for $i = 1, 2$, $0 \ll \frac{K_i(n_v)}{n_v} \leq \frac{A(n_v)}{n_v} \xrightarrow{v \rightarrow \infty} 0$. Given $\epsilon > 0$ and $\delta > 0$ there exists an n' such that for $n_v \geq n'$, f_u has a discrete U-shaped density, by Lemma 3.1, and

$\frac{K_1(n_v)}{n_v} < \epsilon$ and $\frac{K_2(n_v)}{n_v} < \epsilon$, so that

$$\int_0^\epsilon \mathcal{B}_{\frac{W}{\beta}, \frac{B}{\beta}}(p) dp \ll \frac{\delta}{2} \text{ and } \int_{1-\epsilon}^1 \mathcal{B}_{\frac{W}{\beta}, \frac{B}{\beta}}(p) dp \ll \frac{\delta}{2}.$$

We can choose a subset $A_{n_v}^*$, of size $A(n_v)$, from the set of integers $\{1, \dots, n_v\}$, which has maximal probability under single Polya sampling at time n_v . Since f_u is U-shaped for large n , the integer $y \in A_{n_v}^*$ will be a member of one of the sets $\{0, \dots, K_1(n)\}$ or $\{K_2(n), \dots, n\}$.

We want to show that $\sum_{\substack{x_{n_v} \in A_{n_v}^*}} P_u(x_{n_v}) \xrightarrow{v \rightarrow \infty} 0$.

Let $P_u^{n_v}(y)$ be the probability that $\sum_{i=1}^{n_v} X_i = y$ under single Polya sampling at time n_v . Choose $n_v > n'$, then

$$\sum_{\substack{x_{n_v} \in A_{n_v}^*}} P_u(x_{n_v}) \leq \sum_{y \in A_{n_v}^*} P_u^{n_v}(y) = P_u^{n_v}(0 \leq \sum_{i=1}^{n_v} X_i \leq K_1(n_v))$$

$$+ P_u^{n_v}(n_v - K_2(n_v) \leq \sum_{i=1}^{n_v} X_i \leq n_v)$$

$$= \sum_{y=0}^{K_1(n_v)} P_u^{n_v}(y) + \sum_{y=n_v-K_2(n_v)}^{n_v} P_u^{n_v}(y)$$

$$\begin{aligned}
&= \frac{K_1(n_v)}{n_v} \cdot n_v \sum_{y=0}^{n_v} P_u^{n_v}(y) + \sum_{y=n_v}^{n_v} \frac{K_2(n_v)}{n_v} n_v P_u^{n_v}(y) \\
&\leq P_u^{n_v} (0 \leq \sum_{i=1}^{n_v} X_i \leq \varepsilon \cdot n_v) + P_u^{n_v} (n_v(1-\varepsilon) \leq \sum_{i=1}^{n_v} X_i \leq n_v) \\
&\xrightarrow{v \rightarrow \infty} \int_0^\varepsilon \mathcal{B}_{\frac{W}{\beta}, \frac{B}{\beta}}(p) dp + \int_{1-\varepsilon}^1 \mathcal{B}_{\frac{W}{\beta}, \frac{B}{\beta}}(p) dp \leq \delta
\end{aligned}$$

by (3.1).

Case 2: $\frac{W}{\beta} > 1$, $\frac{B}{\beta} > 1$. Choose sequences of integers $\{K_1(n)\}$ and $\{K_2(n)\}$ such that $0 \leq K_1(n)$, $K_2(n) \leq n$, and $A(n) = K_1(n) + K_2(n)$ for all n . Since $\lim_{n \rightarrow \infty} \frac{A(n)}{n} = 0$, there exist a subsequence n_v such that

$$\frac{A(n_v)}{n_v} \xrightarrow{v \rightarrow \infty} 0,$$

then for $i = 1, 2$, $0 \leq \frac{K_i(n_v)}{n_v} \leq \frac{A(n_v)}{n_v} \xrightarrow{v \rightarrow \infty} 0$.

Given $\varepsilon > 0$ and $\delta > 0$, there exists an n' such that for $n_v \geq n'$, f_u is unimodal by Lemma 3.1,

$$\frac{K_1(n_v)}{n_v} < \varepsilon, \quad \frac{K_2(n_v)}{n_v} < \varepsilon$$

and

$$\int_{m-\epsilon}^{m+\epsilon} \mathcal{B}_{\frac{W}{\beta}, \frac{B}{\beta}}(p) dp < \delta,$$

where m is the mode of $\mathcal{B}_{\frac{W}{\beta}, \frac{B}{\beta}}$. Let $n_v > n'$. Since f_u is unimodal,

the set of size $A(n_v)$ with maximal probability with respect to single Polya sampling at time n_v is $A_{n_v}^* = \{m(n_v) - K_1(n_v), m(n_v) - K_1(n_v) + 1, \dots, m(n_v) + K_1(n_v) - 1, m(n_v) + K_1(n_v)\}$, where $m(n_v)$ is the mode of f_u at time n_v . Then

$$\begin{aligned} \sum_{\substack{x \in A_{n_v} \\ x_{n_v}}} P_u(x_{n_v}) &\leq \sum_{y \in A_{n_v}^*} P_u^{n_v}(y) \\ &= P_u^{n_v}(m(n_v) - K_1(n_v) \leq \sum_{i=1}^{n_v} X_i \leq m(n_v) + K_2(n_v)) \\ &= P_u^{n_v}\left(\frac{m(n_v)}{n_v} \cdot n_v - \frac{K_1(n_v)}{n_v} n_v \leq \sum_{i=1}^{n_v} X_i\right) \\ &\leq \frac{m(n_v)}{n_v} n_v + \frac{K_2(n_v)}{n_v} n_v \leq P_u^{n_v}\left(\left(\frac{m(n_v)}{n_v} - \epsilon\right) n_v\right) \\ &\leq \sum_{i=1}^{n_v} X_i \leq \left(\frac{m(n_v)}{n_v} + \epsilon\right) n_v \\ &\xrightarrow{v \rightarrow \infty} \int_{m-\epsilon}^{m+\epsilon} \mathcal{B}_{\frac{W}{\beta}, \frac{B}{\beta}}(p) dp < \delta \end{aligned}$$

by (3.1) and Lemma 3.2.

There is no need to treat the case where $\frac{W}{\beta} = \frac{B}{\beta} = 1$ separately, since by Lemma 3.1, f_u is discrete uniform for all n , so we can choose our set of maximal probability anywhere we wish. We will consider this as part of Case 3.

Case 3: $\frac{W}{\beta} \leq 1, \frac{B}{\beta} \geq 1$. Since $\lim_{n \rightarrow \infty} \frac{A(n)}{n} = 0$, there exist a subsequence n_v such that $\frac{A(n_v)}{n_v} \xrightarrow{v \rightarrow \infty} 0$.

Given $\epsilon > 0$ and $\delta > 0$, there exist an n' such that for $n_v \geq n'$, $\frac{A(n_v)}{n_v} \leq \epsilon$ and $\int_0^\epsilon \mathcal{Z}_{\frac{W}{\beta}, \frac{B}{\beta}}(p) dp \leq \delta$.

Since f_u is monotone non-increasing, the set of size $A(n_v)$ of maximal probability, under single Polya sampling of size n_v , is $A_{n_v}^* = \{0, \dots, A(n_v)\}$. Let $n_v \geq n'$, then

$$\begin{aligned} \sum_{\substack{x_{n_v} \\ x_{n_v} \in A_{n_v}^*}} P_u(x_{n_v}) &\leq \sum_{y \in A_{n_v}^*} P_u^{n_v}(y) = \sum_{y=0}^{A(n_v)} P_u^{n_v}(y) = \frac{A(n_v)}{n_v} \cdot n_v \\ &= \sum_{y=0}^{A(n_v)} P_u^{n_v}(y) \\ &\leq P_u^{n_v}(0 \leq \sum_{i=1}^{n_v} X_i \leq \epsilon \cdot n_v) \xrightarrow{n_v \rightarrow \infty} \int_0^\epsilon \mathcal{Z}_{\frac{W}{\beta}, \frac{B}{\beta}}(p) dp \leq \delta \end{aligned}$$

by (3.1).

Case 4: $\frac{W}{\bar{p}} \geq 1$, $\frac{B}{\bar{p}} \leq 1$, $W \neq B$. Since $\lim_{n \rightarrow \infty} \frac{A(n)}{n} = 0$ there exist a subsequence n_v such that $\frac{A(n_v)}{n_v} \rightarrow 0$. Given $\epsilon > 0$, $\delta > 0$, there exist an n' such that, if $n_v \geq n'$,

$$\frac{A(n_v)}{n_v} \leq \epsilon$$

and

$$\int_{1-\epsilon}^1 \mathcal{B}_{\frac{W}{\bar{p}}, \frac{B}{\bar{p}}}(p) dp \leq \delta.$$

Since f_u is monotone increasing, the set of size $A(n_v)$ of maximal probability, under single Polya sampling at time n_v , is $A_{n_v}^* = \{n_v - A(n_v) + 1, \dots, n_v\}$. Let $n_v > n'$, then

$$\sum_{x_{n_v} \in A_{n_v}} P_u(x_{n_v}) \leq \sum_{y \in A_{n_v}^*} P_u^{n_v}(y) = \sum_{y=n_v-A(n_v)}^{n_v} P_u^{n_v}(y)$$

$$= \sum_{y=n_v(1-\frac{A(n_v)}{n_v})}^{n_v} P_u^{n_v}(y)$$

$$\leq P_u^{n_v}(n_v(1-\epsilon)) \leq \sum_{i=1}^{n_v} X_i \leq n_v$$

$$\xrightarrow{v \rightarrow \infty} \int_{1-\epsilon}^1 \mathcal{B}_{\frac{W}{\bar{p}}, \frac{B}{\bar{p}}}(p) dp < \delta$$

by (3.1).

qed

In the case where the $A(n)$'s are bounded, the hypothesis of Theorems 3.1 is satisfied, leading to the following corollary.

Corollary 3.1 For a stop rule S , if $A(n)$ is bounded, then S is closed.

Proof: Given $A(n)$ bounded, there exist a constant m such that $0 \leq m \leq \infty$, and $A(n) \leq m$ for all n . Then

$$\lim_{n \rightarrow \infty} \frac{A(n)}{n} \leq \lim_{n \rightarrow \infty} \frac{m}{n} = 0.$$

Applying Theorem 3.1, we are finished.

qed

Since the proof of Theorem 3.1 is long and tedious, one may be inclined to try a more direct proof using the uniform convergence of the cumulative distribution function to its limit. It appears, however, that one is left with a requirement on the rate of growth of $A(n)$ in terms of the rate of convergence of the Polya distribution to its beta limit, as is indicated by Theorem 3.2.

First, for each n , we define a right-continuous Polya density, $P_{u_n}^*(z)$, $z \in [0,1]$, from the discrete Polya density, $P_{u_n}(y)$, $y \in \{0, \dots, n\}$, as follows:

$$P_{u_n}^*(z) = \begin{cases} P_{u_n}([nz]) \cdot 2n & \text{for } z \in [\frac{2[nz]-1}{2n}, \frac{2[nz]+1}{2n}] \\ P_{u_n}(0) \cdot n & \text{for } z < \frac{1}{2n} \\ P_{u_n}(n) \cdot n & \text{for } z \geq 1 - \frac{1}{2n}, \end{cases}$$

where, as in Chapter II, $[y]$ denotes the largest integer in y .

Let $F_{u_n}^*(z)$ be the corresponding distribution function, and $F_{\mathcal{B}}(z)$ the cumulative beta distribution function to which it converges, and define

$$d_n = \sup_{z \in [0,1]} |F_{u_n}^*(z) - F_{\mathcal{B}}(z)|,$$

and

$$g_n = \min(n, \frac{1}{d_n}).$$

Theorem 3.2 If, for a stop rule S , $\lim_{n \rightarrow \infty} \frac{A(n)}{g_n} = 0$, then S is closed.

Proof: Let a_j be the j th distinct value of $\sum_{i=1}^n x_i$ for $\underline{x}_n \in A_n$, where A_n is defined in the proof of Theorem 3.1. Place an interval L_{nj} of radius $\frac{1}{2n}$ about the point $\frac{a_j}{n}$; that is $L_{nj} = [\frac{a_j}{n} - \frac{1}{2n}, \frac{a_j}{n} + \frac{1}{2n}]$, where $L_{nj} = [0, \frac{1}{2n}]$ or $(1 - \frac{1}{2n}, 1)$ if $a_j = 0$ or n , and let $D_n = \bigcup_{i=1}^{A(n)} L_{ni}$.

Denote by $\mu_L(D)$ the Lebesgue measure of the set D , and let $f_{\mathcal{B}}(\cdot)$ be the beta density. By Proposition 13, page 85 of Royden (1968), given $\epsilon > 0$, there exists a $\delta(\epsilon) > 0$ such that

$$\mu_L(D) \leq 2 \delta(\epsilon)$$

implies that

$$\int_D f_{\mathcal{B}}(z) d_z \leq \epsilon. \quad (3.2a)$$

Let $\delta^*(\epsilon) = \min(\epsilon, \delta(\epsilon))$. By assumption, given $\delta^*(\epsilon) > 0$, where exists a $v(\epsilon)$ such that $v > v(\epsilon)$ implies that

$$\frac{A(n_v)}{g(n_v)} \leq \delta^*(\epsilon) \leq \epsilon. \quad (3.2b)$$

Choose n_v such that $v > v(\epsilon)$, then

$$P_{u_{n_v}}(A_{n_v}) = P_{u_{n_v}}^*(D_{n_v}) \leq P_{\mathcal{B}}(D_{n_v}) + |P_{u_{n_v}}^*(D_{n_v}) - P_{\mathcal{B}}(D_{n_v})|,$$

where $P_{\mathcal{B}}(D_{n_v})$ is the probability of D_{n_v} under the limiting beta distribution. Now

$$u_L(D_{n_v}) \leq \frac{2A(n_v)}{n_v} \leq \frac{2A(n_v)}{g_{n_v}} \leq 2\epsilon \quad (3.3)$$

by (3.2b). Since

$$P_{\mathcal{B}}(D_{n_v}) = \int_{D_{n_v}} f_{\mathcal{B}}(z) d u_L(z),$$

$$P_{\mathfrak{F}}(D_{n_v}) < 2\epsilon, \text{ by (3.2a).}$$

It remains only to show that

$$\left| P_{u_{n_v}}^*(D_{n_v}) - P_{\mathfrak{F}}(D_{n_v}) \right|$$

is small,

$$\left| P_{u_{n_v}}^*(D_{n_v}) - P_{\mathfrak{F}}(D_{n_v}) \right|$$

$$\begin{aligned} &\leq \sum_{j=1}^{A(n_v)} \left| P_{u_{n_v}}^*(L_{n_v j}) - P_{\mathfrak{F}}(L_{n_v j}) \right| \\ &= \sum_{j=1}^{A(n_v)} \left| F_{u_{n_v}}^*\left(\frac{a_{n_v j}}{n_v}\right) + \frac{1}{2n_v} - F_{u_{n_v}}^*\left(\frac{a_{n_v j}}{n_v}\right) - \frac{1}{2n_v} \right. \\ &\quad \left. - F_{\mathfrak{F}}\left(\frac{a_{n_v j}}{n_v} + \frac{1}{2n_v}\right) + F_{\mathfrak{F}}\left(\frac{a_{n_v j}}{n_v} - \frac{1}{2n_v}\right) \right| \\ &\leq \sum_{j=1}^{A(n_v)} \left[\left| F_{u_{n_v}}^*\left(\frac{a_{n_v j}}{n_v} + \frac{1}{2n_v}\right) - F_{\mathfrak{F}}\left(\frac{a_{n_v j}}{n_v} + \frac{1}{2n_v}\right) \right| \right. \\ &\quad \left. + \left| F_{u_{n_v}}^*\left(\frac{a_{n_v j}}{n_v} - \frac{1}{2n_v}\right) - F_{\mathfrak{F}}\left(\frac{a_{n_v j}}{n_v} - \frac{1}{2n_v}\right) \right| \right] \end{aligned}$$

$$\begin{aligned}
& A(n_v) \\
& \leq \sum_{j=1}^{\infty} 2d_{n_v}^j, \\
& = 2 A(n_v) d_{n_v}, \\
& \leq 2 \frac{A(n_v)}{g_{n_v}} \\
& < 2 \epsilon \text{ by (3.2).}
\end{aligned}$$

qed

We next investigate the relationships between closure of a stop rule under Bernoulli sampling and closure under Polya sampling.

From the sufficient conditions for closure of a stop rule under Bernoulli and Polya sampling, it is obvious that, for under Bernoulli sampling, $\lim_{n \rightarrow \infty} \frac{A(n)}{\sqrt{n}} < \infty$, then $\lim_{n \rightarrow \infty} \frac{A(n)}{n} = 0$, so that closure holds under Polya sampling as well. This fact, however, does not imply that a stop rule which is closed under Bernoulli sampling will be closed under Polya sampling, since the given condition for closure is a sufficient, but not a necessary one. We will now prove that closure under binomial sampling implies closure under Polya sampling.

Theorem 3.3 If a stop rule S is closed under Bernoulli sampling for all $p \in (0,1)$, then it is closed under Polya sampling for all positive integers W , B , and β .

Proof: Recall the set K_n , $n=1, 2, \dots$, introduced in the proof of Theorem 3.1, which is the set of path segments in S stopped by time n . Since the K_n 's are a monotonically increasing sequence of sets converging to S , the probabilities, $P_b(K_n)$ and $P_u(K_n)$ are monotonically increasing functions bounded by one. Therefore $\lim_{n \rightarrow \infty} P_b(K_n) = P_b(S)$ and $\lim_{n \rightarrow \infty} P_u(K_n) = P_u(S)$. Fix W , B , and β ; then in view of (2.2.5), the Monotone Convergence Theorem (see Theorem 12, page 227 of Royden [1968]), and closure for Bernoulli sampling,

$$\begin{aligned} P_u(S) &= \lim_{n \rightarrow \infty} P_u(K_n) = \int_0^1 \lim_{n \rightarrow \infty} P_b(K_n) \mathcal{B}_{\frac{W}{\beta}, \frac{B}{\beta}}(p) dp \\ &= \int_0^1 P_b(S) \mathcal{B}_{\frac{W}{\beta}, \frac{B}{\beta}}(p) dp = 1. \end{aligned}$$

Since this argument holds for all values of W , B , and β , $P_u(S) = 1$ for all W , B , and β .

qed

It is apparent that when Girshick, Mosteller, Savage and Wolfowitz talked about closure, they meant $P_b(S) = 1$ for all possible values of the parameter p and this has been our view for Polya sampling plans until now.

A less restrictive definition of closure would allow a stop rule to be open ($P(S) < 1$) for certain values of the parameters.

There is some precedence for this in the area of sequential decision rules in the work of Robbins (1970). The fact that the SPRT for Polya sampling is not closed lends some credence to the idea.

Since sampling is undertaken to estimate and test hypotheses about parameters, and in most practical situations, some knowledge of the parameters usually exists, a stop rule which is closed for an appropriate subset of parameters could be employed. This could be desirable if it possesses a smaller expected sample size or an estimate with a smaller variance than stop rules which are closed for all parameter values.

In Theorem 3.4, exploiting the relationship between the Bernoulli and Polya processes further, we show that any stop rule which is open for an interval of p values under Bernoulli sampling will be open for all values of the parameters under Polya sampling.

Theorem 3.4 If a stop rule S is open for an interval of p values under Bernoulli sampling, it is open under Polya sampling for all values of W , B , and β .

Proof: Assume S is open under Bernoulli sampling for an interval $I \subset (0,1)$; that is

$$P_b(S) \leq 1$$

for all $p \in I$.

Recall the sets A_n and K_n , $n=1, 2, \dots$, defined in the proof of Theorem 3.1, which are, respectively, the set of all path segments

for which sampling continues at time n , and the set of all path segments for which sampling has stopped by time n . Since

$$1 = P_b(A_n) + P_b(K_n),$$

and, $P_b(K_n)$ and $P_u(K_n)$ are monotonically increasing functions, $P_b(A_n)$ and $P_u(A_n)$ must be decreasing functions which are bounded below by zero. Therefore, the limits,

$$\lim_{n \rightarrow \infty} P_b(A_n) \quad \text{and} \quad \lim_{n \rightarrow \infty} P_u(A_n)$$

exist, and for Bernoulli sampling,

$$\lim_{n \rightarrow \infty} P_b(A_n) = 1 - P_b(S) > 0. \quad (3.4)$$

Fix W , B , and β ; by the Lebesgue Convergence Theorem (see Theorem 16, page 229 of Royden [1968]), and (3.4),

$$\begin{aligned} \lim_{n \rightarrow \infty} P_u(A_n) &= \int_0^1 \lim_{n \rightarrow \infty} P_b(A_n) \mathcal{B}_{\frac{W}{\beta}, \frac{B}{\beta}}(p) dp \\ &= \int_0^1 (1 - P_b(S)) \mathcal{B}_{\frac{W}{\beta}, \frac{B}{\beta}}(p) dp \\ &> 0. \end{aligned}$$

Since this argument holds for all W , B , and β ,

$$P_u(S) < 1$$

for all W , B , and β .

qed

We now turn our attention to GMS stop rule partitions. Recall that we have defined the closure of stop rules, that is, $P(S) = 1$. However, when dealing with GMS stop rule partitions, it seems natural to think of the boundaries as being closed, and therefore the stop rule partition S_π being closed.

In Theorem 3.5, we establish that if a GMS stop rule partition is described by a boundary approximated by two diverging straight lines, then it is open under Polya sampling for all parameter values.

Theorem 3.5 Consider a GMS stop rule partition S_π , whose sampling region R lies between the points (n, d_{1n}) and (n, d_{2n}) for each n . Suppose that (i) there exist constants s_1 and s_2 , $0 < s_1 < s_2 \leq 1$, such that $\lim_{n \rightarrow \infty} \frac{d_{2n}}{n} \geq s_2$, and $\overline{\lim}_{n \rightarrow \infty} \frac{d_{1n}}{n} \leq s_1$, and (ii) the sequence of points $\{(n, \lceil n(s_1 + s_2)/2 \rceil)\}$, where $\lceil y \rceil$ is the greatest integer in y , belong to R for all n . Then S_π is open under Polya sampling for all W , B , and β .

Proof: We will define a subset of R for which, under Bernoulli sampling for some interval of p , there is positive probability that sampling may never stop.

Set $\epsilon > 0$. Choose $n(\epsilon)$ large enough such that, for $n \geq n(\epsilon)$

$$d_{1n} \leq (s_1 + \epsilon) n \quad \text{and} \quad d_{2n} \geq (s_2 - \epsilon) n. \quad (3.5)$$

Suppose that Y_i , $i=1, 2, \dots$, are independent Bernoulli (p) random variables. Let

$$I = \left[\frac{s_1 + s_2}{2} - \epsilon, \frac{s_1 + s_2}{2} + \epsilon \right],$$

then, again, where $[y]$ is the greatest integer in y ,

$$P_b \{ d_{1n} \leq \sum_{i=1}^n Y_i \leq d_{2n} \text{ for all } n, \text{ and } p \in I \}$$

$$\geq P_b \left\{ \sum_{i=1}^n Y_i = \left[n(s_1 + s_2)/2 \right], 1 \leq n \leq n(\epsilon) \right.$$

$$\left. \text{and } d_{1n} \leq \sum_{i=1}^n Y_i \leq d_{2n}, n \geq n(\epsilon) + 1, p \in I \right\}$$

$$= P_b \left\{ \sum_{i=1}^n Y_i = \left[n(s_1 + s_2)/2 \right], 1 \leq n \leq n(\epsilon), p \in I \right\}$$

$$\cdot P_b \{ d_{1r+n(\epsilon)} - \left[n(\epsilon)(s_1 + s_2)/2 \right] \leq \sum_{j=1}^r Y_j \leq$$

$$d_{2r+n(\epsilon)} - \left[n(\epsilon)(s_1 + s_2)/2 \right], r \geq 1, p \in I \}$$

which, by defining

$$\delta(\epsilon) = P_b \left\{ \sum_{i=1}^n Y_i = \lfloor n(s_1+s_2)/2 \rfloor, 1 \leq n \leq n(\epsilon), p \in I \right\},$$

by (3.5), and by the fact that $\lfloor n(\epsilon)(s_1+s_2)/2 \rfloor \leq n(\epsilon)(s_1+s_2)/2$, is greater than

$$\begin{aligned} \delta(\epsilon) P_b \left\{ (s_1+\epsilon)(r+n(\epsilon)) - n(\epsilon)(s_1+s_2)/2 \leq \sum_{j=1}^r Y_j \right. \\ \left. \leq (s_2-\epsilon)(r+n(\epsilon)) - n(\epsilon)((s_1+s_2)/2), r \geq 1, p \in I \right\} \end{aligned}$$

which, by rearranging terms, equals

$$\begin{aligned} \delta(\epsilon) &= P_b \left\{ n(\epsilon) (s_1 + \epsilon - s_1/2 - s_2/2) + r(s_1 + \epsilon) \right. \\ &\leq \sum_{j=1}^r Y_j \leq n(\epsilon) (s_2 - \epsilon - s_1/2 - s_2/2) + r(s_2 - \epsilon), \\ &\geq 1, p \in I \} \\ &= \delta(\epsilon) P_b \left\{ n(\epsilon)((s_1-s_2)/2 + \epsilon) + r(s_1 + \epsilon) \right. \\ &\leq \sum_{j=1}^r Y_j \leq n(\epsilon)((s_2-s_1)/2 - \epsilon) + r(s_2-\epsilon), \\ &\geq 1, p \in I \}. \end{aligned} \tag{3.6}$$

Now choose $a(\epsilon) > 0$ and $b(\epsilon) > 0$ such that

$$a(\epsilon) b(\epsilon) > 1, \quad (3.7)$$

$$\begin{aligned} n(\epsilon) &= ((s_1 - s_2)/2 + \epsilon) + \nu(s_1 + \epsilon) \\ &\leq -b(\epsilon) + \nu(p - a(\epsilon) pq), \end{aligned} \quad (3.8)$$

for $\nu \geq 1$ and $p \in I$,

$$\begin{aligned} n(\epsilon) &= ((s_2 - s_1)/2 - \epsilon) + \nu(s_2 - \epsilon) \\ &\geq b(\epsilon) + \nu(p + a(\epsilon) pq), \end{aligned} \quad (3.9)$$

for $\nu \geq 1$ and $p \in I$.

In view of (3.8) and (3.9), (3.6) is larger than

$$\begin{aligned} \delta(\epsilon) P_b \{ \nu p - \nu a(\epsilon) pq - b(\epsilon) \leq \sum_{j=1}^{\nu} Y_i \leq \nu p + \nu a(\epsilon) pq \\ + b(\epsilon), \nu \geq 1, p \in I \} \end{aligned}$$

which, by Corollary (7) of Dubins and Freedman (1965), is greater than

$$\begin{aligned} \delta(\epsilon) \frac{a(\epsilon) b(\epsilon) - 1}{a(\epsilon) b(\epsilon) + 1} \\ > 0. \end{aligned} \quad (3.10)$$

for $p \in I$, by (ii) and (3.7). By Theorem 3.4, S is therefore open under Polya sampling.

qed

An important class of stop rules are those associated with the Sequential Probability Ratio Test (SPRT). Wald (1947) derived the binomial SPRT for the hypothesis $H_0: p = p_1$ vs $H_1: p = p_0$; and showed that the sampling region, which consists of the points (n, d) lying between two parallel straight lines, is closed for all $p \in [0, 1]$.

It will be shown in Lemma 3.6 that the analogous SPRT sampling region in the Polya case lies between two diverging lines, and is open for all values of the parameters W , B , and β . Lemmas 3.3-3.5 establish some necessary preliminaries.

Fix $N = W + B$, β , and n and let $W_1 > W_0$. Define

$$f_n(p) = \frac{P_u(\lceil np \rceil, n; W_1, B_1, \beta)}{P_u(\lceil np \rceil, n; W_0, B_0, \beta)} = \frac{\prod_{i=0}^{\lceil np \rceil - 1} (W_1 + i\beta) \prod_{i=0}^{n - \lceil np \rceil - 1} (B_1 + i\beta)}{\prod_{i=0}^{\lceil np \rceil - 1} (W_0 + i\beta) \prod_{i=0}^{n - \lceil np \rceil - 1} (B_0 + i\beta)} \quad (3.11)$$

for $p \in (0, 1)$, where $\lceil y \rceil$ is the largest integer in y . Note that if np is an integer, say d , $f_n(\frac{d}{n})$ is the likelihood ratio

$$\frac{P_u(d, n; W_1, B_1, \beta)}{P_u(d, n; W_0, B_0, \beta)}.$$

Lemma 3.3 For fixed N , β , and n , $f_n(p)$ is non-decreasing in $p \in [0,1]$.

Proof: Let $W_1 > W_0$ and $p_1 > p_0$, then $B_0 > B_1$ and for all $i, j=0, 1, \dots$,

$$\frac{W_1 + i\beta}{W_0 + i\beta} \geq \frac{B_1 + j\beta}{B_0 + j\beta};$$

therefore

$$\frac{\prod_{i=0}^{[np_1]-1} (W_1 + i\beta)}{\prod_{i=0}^{[np_1]-1} (W_0 + i\beta)} \frac{\prod_{i=0}^{n-[np_1]-1} (B_1 + i\beta)}{\prod_{i=0}^{n-[np_1]-1} (B_0 + i\beta)} \geq \frac{\prod_{i=0}^{[np_0]-1} (W_1 + i\beta)}{\prod_{i=0}^{[np_0]-1} (W_0 + i\beta)} \frac{\prod_{i=0}^{n-[np_0]-1} (B_1 + i\beta)}{\prod_{i=0}^{n-[np_0]-1} (B_0 + i\beta)}.$$

qed

Lemma 3.4 For fixed N , β , and $W_1 > W_0$, $f_n(p)$ converges uniformly to

$$f(p) = \frac{\Gamma\left(\frac{W_0}{\beta}\right) \Gamma\left(\frac{B_0}{\beta}\right)}{\Gamma\left(\frac{W_1}{\beta}\right) \Gamma\left(\frac{B_1}{\beta}\right)} \left(\frac{p}{1-p}\right)^{\frac{W_1 - W_0}{\beta}} \quad (3.12)$$

for $p \in [a,b] \subset (0,1)$.

Proof: Utilizing (12.25), p. 66 of Feller (1968), in (3.11) we have

$$\begin{aligned}
 f_n(p) &= \frac{\Gamma(\frac{W_0}{\beta}) \Gamma(\frac{B_0}{\beta})}{\Gamma(\frac{W_1}{\beta}) \Gamma(\frac{B_1}{\beta})} |np|^{(\frac{W_1-W_0}{\beta})} (n-|np|)^{(\frac{B_1-B_0}{\beta})} + o(1) \\
 &= \frac{\Gamma(\frac{W_0}{\beta}) \Gamma(\frac{B_0}{\beta})}{\Gamma(\frac{W_1}{\beta}) \Gamma(\frac{B_1}{\beta})} \left(\frac{|np|}{n-|np|}\right)^{\frac{W_1-W_0}{\beta}} + o(1),
 \end{aligned}$$

since $W_1 - W_0 = B_0 - B_1$.

Therefore

$$f(p) = \lim_{n \rightarrow \infty} f_n(p) = \frac{\Gamma(\frac{W_0}{\beta}) \Gamma(\frac{B_0}{\beta})}{\Gamma(\frac{W_1}{\beta}) \Gamma(\frac{B_1}{\beta})} \left(\frac{p}{1-p}\right)^{\frac{W_1-W_0}{\beta}}$$

for $p \in (0,1)$.

The uniformity of this convergence now follows by a standard argument:

Establish a grid of fineness, ϵ , on $[f(a), f(b)]$, and, by application of the monotone function $f^{-1}(\cdot)$, a corresponding grid on $[a,b]$. Denote the points of that grid by

$$a = a_0 < a_1 \dots < a_k = b.$$

Since there are a finite number of a_i 's, choose $n(\epsilon)$ such that

$$|f_n(a_i) - f(a_i)| < \epsilon \quad (3.13)$$

for $i=0, \dots, k$ and $n > n(\epsilon)$. Let $t \in (a_i, a_{i+1})$ and $n > n(\epsilon)$, then

$$\begin{aligned} |f_n(t) - f(t)| &\leq |f_n(t) - f(a_i)| + |f(a_i) - f(a_{i+1})| \\ &\quad + |f(a_{i+1}) - f(t)|. \end{aligned}$$

Now $|f(a_i) - f(a_{i+1})| \leq \epsilon$ and $|f(a_{i+1}) - f(t)| < \epsilon$ by construction, and

$$\begin{aligned} |f_n(t) - f(a_i)| &< |f_n(a_{i+1}) - f(a_i)| \leq |f_n(a_{i+1}) - f(a_{i+1})| \\ &\quad + |f(a_{i+1}) - f(a_i)| \leq 2\epsilon, \end{aligned}$$

by construction and the convergence of $f_n(\cdot)$ to $f(\cdot)$.

qed

Recall that for constants, $0 < B < 1 \leq A$, chosen by the desired levels of the Types I and II errors, α and β , respectively, the SPRT is defined as follows:

$$\text{accept } H_1 \text{ if } f_n\left(\frac{d}{n}\right) \geq A,$$

accept H_0 if $f_n\left(\frac{d}{n}\right) \leq B$,

and

continue sampling if $B \leq f_n\left(\frac{d}{n}\right) \leq A$.

The boundary of the Polya SPRT, in view of Lemma 3.3, for every n , consists of the points (n, d_{1n}') and (n, d_{2n}') where

$$d_{1n}' = \max_d \left\{ d: f_n\left(\frac{d}{n}\right) \leq B \right\}$$

and

$$d_{2n}' = \min_d \left\{ d: f_n\left(\frac{d}{n}\right) \leq A \right\}.$$

Lemma 3.5 now shows that the boundary of the Polya SPRT is "essentially" given by two diverging straight lines.

Lemma 3.5 For a Polya SPRT with critical values $B \leq A$ and boundary consisting of the points (n, d_{1n}') and (n, d_{2n}') for each n , $\frac{d_{1n}'}{n} \xrightarrow{n \rightarrow \infty} f^{-1}(B)$ and $\frac{d_{2n}'}{n} \xrightarrow{n \rightarrow \infty} f^{-1}(A)$, where f is defined in (3.12).

Proof: Taking $a \leq f^{-1}(B)$ and $b > f^{-1}(A)$ in Lemma 3.4, it is clear

that $\frac{d_{1n}'}{n} \xrightarrow{n \rightarrow \infty} f^{-1}(B)$ and $\frac{d_{2n}'}{n} \xrightarrow{n \rightarrow \infty} f^{-1}(A)$.

qed

We are now in a position to prove Polya SPRT's are not closed.

Lemma 3.6 Polya SPRT sampling regions are open for all values of the parameters W , B , β .

Proof: Apply Theorem 3.5 with $s_1 = f^{-1}(B)$ and $s_2 = f^{-1}(A)$.

qed

Even though the Polya SPRT is not closed, bounds on the Types I and II errors may be calculated in a manner, similar to the well-known method for the closed case. Define the sets

$$D_1 = \{(n, d'_{1n})\}$$

and

$$D_2 = \{(n, d'_{2n})\},$$

that is, the set of points for which H_0 is accepted, and the set of points for which H_1 is accepted, respectively.

For critical constants $B < A$,

$$\frac{P_u(D_2; N, W_1, \beta)}{P_u(D_2; N, W_0, \beta)} \geq A, \quad (3.14)$$

$$\frac{P_u(D_1 | N, W_1, \beta)}{P_u(D_1 | N, W_0, \beta)} \leq B. \quad (3.15)$$

The lack of closure implies

$$P_u\{D_2; N, W_1, \beta\} + P_u\{D_1; N, W_1, \beta\} = \eta \leq 1$$

and

$$P_u\{D_2; N, W_0, \beta\} + P_u\{D_1; N, W_0, \beta\} = \delta \leq 1.$$

Solving for $P_u\{D_u; N, W_1, \beta\}$ and $P_u(D_1; N, W_0, \beta)$, respectively, and substituting in (3.14) and (3.15), where $\varphi = P_u\{D_2; N, W_0, \beta\}$ and $\lambda = P_u\{D_1; N, W_1, \beta\}$, η and δ satisfy the following inequalities.

$$\frac{\eta - \lambda}{\varphi} \geq A$$

and

$$\frac{\lambda}{\delta - \varphi} \leq B.$$

This of course would only be useful if η or δ are either known or estimable.

IV. SIMPLICITY, DECOMPOSABILITY, AND THEIR CONSEQUENCES

For every positive integer n , let D_n be a subset of the non-negative integers $(0, 1, \dots, n)$. The stop rules considered by Girshick, Mosteller, and Savage (1946), Wolfowitz (1946), and Savage (1947) involve a stop rule partition S_π based on a sequence (D_1, D_2, \dots) in the following way: If z_i is the i^{th} Bernoulli sample, sampling stops the first time that $\sum_{i=1}^n z_i \in D_n$. If $D_n = \{d_{n1}, \dots, d_{nk_n}\}$, then it is natural (as indeed was done by the aforementioned authors) to call the set of points,

$$\{(1, d_{11}), \dots, (1, d_{1k_1}); (2, d_{21}), \dots, (2, d_{2k_2}) \dots\},$$

the boundary.

The geometrically suggestive term "boundary" is appropriate for this reason: If the progress of sampling is documented by plotting the successive values of $(n, \sum_{i=1}^n z_i)$ in the plane, the stop rule is properly portrayed by literally construing the points (i, d_{ij}) as obstacles to the progress of the "path" $\{(1, z_1), (2, z_1+z_2), \dots\}$.

Thus, the GMS stop rule partitions studied by Girshick, Mosteller, Savage (1946), Wolfowitz (1946), and Savage (1947) are representable geometrically, and it is not totally surprising that the entirely geometric property of simplicity, to be defined below, plays an important role in their treatment of completeness.

In contrast to the GMS stop rule partitions that Girshick, Mosteller, Savage (1946), Wolfowitz (1946), and Savage (1947) studied, the discussion in Chapter V treats more general stop rule partitions, which, rather than being themselves boundaries, induce certain "derived point configurations" in the (n, d) plane. Hence our terminology intentionally is made to differ from that of the above four authors in that it distinguishes stop rule partition from boundary, as well as the relevant properties associated with each.

Consider a finite sequence of points in the plane $\{(x_i, y_i)\}_{i=1}^n$ with the following properties:

- (i) $(x_0, y_0) = (0, 0)$,
- (ii) x_i and y_i are non-negative integers such that $y_i \leq x_i$
- (iii) either
 - (iiia) $(x_{i+1}, y_{i+1}) = (x_i + 1, y_i)$
 - or
 - (iiib) $(x_{i+1}, y_{i+1}) = (x_i + 1, y_i + 1)$.

Such a sequence of points will be called a path segment, consisting of horizontal steps (iiia) and diagonal steps (iiib), starting at the origin, with path stop point (x_n, y_n) .

Remark 4.1 Note that, in Chapter I, we represented a path segment as an n -vector of zeros and ones. Now we have represented it as a finite sequence of points, $\{(x_i, y_i)\}_{i=1}^n$. It should be obvious that the two representations are equivalent: we can derive the sequence of points from the vector, and conversely.

Denote by \mathcal{P} the points of the plane with non-negative integer coordinates (x, y) satisfying $x \geq y$. Let ζ be a configuration of points in \mathcal{P} , and a point of ζ be called a ζ -point. ζ partitions the points of \mathcal{P} into three types given by the following three definitions:

Definition 4.1 A point of \mathcal{P} is origin-disconnected (O-D) if every path from the origin to the point passes through at least one ζ -point.

Definition 4.2 A point of \mathcal{P} is a stop point if it belongs to ζ and is not an O-D point.

Definition 4.3 A point of \mathcal{P} is origin-connected (O-C) if it is neither a stop nor an O-D point.

In addition, we give the following definitions:

Definition 4.4 The set of all stop points in ζ is called the lining of ζ , denoted by $L(\zeta)$.

Definition 4.5 (Girshick, Mosteller and Savage [1946]). ζ is simple if each vertical segment between two O-C points contains only O-C points.

David and Olkin (1956) also investigated GMS stop rule partitions, but in the case of geometric sampling. They proposed another geometric property, in their study of completeness, namely decomposability, which will be defined below.

Definition 4.6 (David and Olkin [1956]) A horizontal (diagonal) line $y = b$ ($y = x - a$), $a, b = 0, 1, 2, \dots$, is called a b-cut (a-cut).

Definition 4.7 (David and Olkin [1956]). ζ is decomposable if there exist a sequence of cuts with the following properties:

- (i) The subsequence of b-cuts (a-cuts) is $y = 0, y = 1, y = 2, \dots$
 $(y = x, y = x - 1, y = x - 2, \dots)$.
- (ii) Every cut of the sequence contains exactly one point of $L(\zeta)$ not contained in a previous cut;
- (iii) Every point of $L(\zeta)$ lies on some cut of the sequence.

Another geometric property is the following:

Definition 4.8 (see Sobel [1953]) If there exist either or both (possibly finite) double sequences $\{(n_1, c_1), (n_2, c_2), \dots\}$ and $\{(n'_1, d_1), (n'_2, d_2), \dots\}$ for which $n_i \leq n_{i+1}$, $c_i \leq c_{i+1}$, $n'_i \leq n'_{i+1}$, $d_{i+1} < d_i$, and $c_i < d_i$, and if $L(\zeta)$ contains either or both sets of points

$$\begin{aligned}
L = \{ & (n_1, 0), (n_1, 1), \dots, (n_1, c_1), \\
& (n_2, c_1+1), (n_2, c_1+2), \dots, (n_2, c_2), \\
& (n_3, c_2+1), (n_3, c_2+2), \dots, (n_3, c_3), \\
& \dots \dots \dots \dots \dots \dots \}
\end{aligned}$$

or

$$\begin{aligned}
U = \{ & (n_1', d_1), (n_1', d_1+1), \dots, (n_1', n_1'), (n_2', d_2), (n_2', d_2+1), \\
& \dots, (n_2', d_1-1+n_2'-n_1'), (n_3', d_3'), (n_3', d_3'+1), \dots, \\
& (n_3', d_2-1+n_3'-n_2'), \dots \} ,
\end{aligned}$$

then ζ is of the GSPRT-type (for Generalized Sequential Probability Ratio Test).

Note that if $n_i = n_i'$, the well-known multiple sampling scheme is a special case of the GSPRT-type.

Remark 4.2 It should be mentioned that our definitions of simplicity, decomposability, and GSPRT-type are stated as properties of the point configuration, ζ , while their conditions are restrictions on the lining $L(\zeta)$. The reason for this distinction will become apparent in the next chapter.

The relationship between simplicity and decomposability is given in the following theorem.

Theorem 4.1 If ζ is simple, it is decomposable.

Proof: Assume ζ is simple. Let m_1 be the smallest of the first coordinates of the points of $L(\zeta)$. All the points lying on the line $x = m_1$ must be O-C or stop points. By simplicity, the O-C points must be contiguous, which means that $(m_1, 0)$ and (m_1, m_1) cannot both be O-C points.

Assume, without loss of generality, that (m_1, m_1) is a stop point; then it lies on the cut $y - x = 0$, which, by the definition of stop point, can contain no other stop points. Hence, $x - y = 0$ serves as the first cut.

Suppose that there is another stop point on the line $x = m_1$. Recall from the above, that all points on the line $x = m_1$ are either O-C or stop points. Then, as before, $(m_1, m_1 - 1)$ and $(m_1, 0)$ cannot both be O-C points, in view of simplicity. Suppose, without loss of generality, that $(m_1, m_1 - 1)$ is a stop point, then it lies on the cut $x - y = 1$, and there cannot be another stop point on the cut $x - y = 1$. Hence $x - y = 1$ serves as the second cut.

The stop point(s) on $x = m_1$ thus provides (provide) the initial cut(s), and, if all the points on $x = m_1$ are stop points, in fact provide all the cuts.

Suppose that not all the stop points lie on $x = m_1$. Let m_2 be the smallest of the first coordinates of the stop points not on $x = m_1$.

The cuts, say $y = 0, y = 1, \dots, y = a, x - y = 0, x - y = 1,$

... , $x - y = b$, $a, b \geq 0$, provided by the stop points on $x = m_1$ yield two possibly contiguous sets of O-D points on the line $x = m_2$:

$$\{(m_2, 0), \dots, (m_2, a)\}$$

and

$$\{(m_2, m_2 - b), \dots, (m_2, m_2)\} ,$$

and the remaining contiguous set of points $\{(m_2, a+1), (m_2, a+2), \dots, (m_2, m_2 - n - 1)\}$ must contain only O-C and stop points. As before, both $(m_2, a+1)$ and $(m_2, m_2 - b - 1)$ cannot both be O-C points, since simplicity would then be violated. Suppose, therefore, for example, that $(m_2, a+1)$ is a stop point, it lies on the cut $y = a+1$, and, since, in view of the stop points uncovered so far, there can be no other stop point on this cut, so $y = a+1$ serves as the next cut.

The remaining stop points on $x = m_2$ provide cuts in a manner analogous to the manner in which cuts were provided by the points with first coordinate m_1 .

If there are remaining stop points not on $x = m_2$, define m_3 as the smallest of the first coordinates of these remaining stop points, and proceed to find cuts as was done for the stop points with first coordinate m_1 and first coordinate m_2 .

The process continues until all stop points are matched with cuts.

qed

If one assumes the closure of ζ (in which case almost all path segments are stopped by points of $L(\zeta)$), then decomposability implies simplicity and GSPRT-type. Since GSPRT-type is obviously simple, it is possible to establish an equivalence among the three properties.

If ζ is also finite, the equivalence can be carried one step further.

Definition 4.9 ζ is finite if there is an integer m such that no stop point of ζ has first coordinate greater than m .

Plackett (1948) showed that a closed, finite boundary (GMS stop rule partition) which has exactly $m+1$ stop points, where m is the maximum sample size, is simple.

We can establish, using Plackett's (1948) result and Theorem 4.1, the following

Theorem 4.2 If ζ is closed and finite, then (a) ζ is simple \Leftrightarrow (b) ζ is decomposable \Leftrightarrow (c) ζ is of GSPRT-type \Leftrightarrow (d) ζ contains exactly $m+1$ boundary points.

Proof: a) \Leftrightarrow b) by Theorem 4.1.

b) \Leftrightarrow c): Let m_1 be the smallest of the first coordinates of the points of $L(\zeta)$. Assume there are $j_1 \geq 1$ stop points on the line $x = m_1$, then by closure and decomposability, there is a set of j_1 diagonal and vertical cuts

$$x - y = 0, \dots, x - y = a,$$

and

$$y = 0, \dots, y = j_1 - 2 - a,$$

$a \geq 0$, associated with the respective stop points:

$$(m_1, m_1), \dots, (m_1, m_1 - a),$$

and

$$(m_1, 0), \dots, (m_1, j_1 - 2 - a).$$

So for $n_1 = m_1$, $c_1 = j_1 - 2 - a$ and $d_1 = m_1 - a$.

Let m_2 be the smallest of the first coordinates of the remaining points of $L(\zeta)$. Suppose there are $j_2 \geq 1$ stop points on the line $x = m_2$. The points $(m_2, 0), \dots, (m_2, j_1 - 2 - a)$ and $(m_2, m_2), \dots, (m_2, m_2 - a)$ are O-D points in view of the cuts provided by the stop points on $x = m_1$. The cuts provided by the stop points on $x = m_2$ are

$$x - y = a + 1, \dots, x - y = a + 1 + b$$

and

$$y = j_1 - 1 + a, \dots, y = j_1 - 1 + a + j_2 - 2 - b$$

$b \geq 0$, corresponding to the stop points

$$(m_2, m_2 - a - 1), \dots, (m_2, m_2 - a - 1 - b)$$

and

$$(m_2, j_1 - 1 + a), \dots, (m_2, j_1 + j_2 - 3 + a - b).$$

So that for $n_2 = m_2$, $c_2 = j_1 + j_2 - 3 + a - b$ and $d_2 = m_2 - a - 1 - b$.

In the same manner, c_i and d_i may be determined successively for each line $x = n_i$ containing stop points.

c) \Rightarrow a) Obvious.

d) \Rightarrow a) By Plackett (1948).

b) \Rightarrow d): Let k_i be the number of stop points on the line $x = m_i$. For $x = m$, the last line which contains stop points, there are at most $m+1$ points in \mathcal{P} . By closure and decomposability, the points on $x = m$ lying on the cuts provided by the k_i stop points, $i < m$, are all O-D points. The number of O-C points on $x = m$ must then be

$$m+1 - \sum_{i \leq m} k_i;$$

but since ζ is closed at $x=m$, this implies

$$m+1 - \sum_{i \leq m} k_i = 0.$$

Therefore, the number of stop points must be $m+1$.

qed

V. COMPLETENESS

The purpose of this chapter is to investigate completeness in the context of Polya sampling for stop rules partitions which are not necessarily GMS.

Two examples of stop rule partitions which are not GMS are

- (i) Stop after three consecutive white draws, and
- (ii) Stop the first time

$$\sum_{i=1}^n c_i x_i = k_n$$

where the c_i are not identically one.

We now introduce notation for stop rule partitions that are not GMS:

Let \mathcal{X}' denote the space of infinite sequences (x_1, x_2, \dots) with $x_i = 0$ or 1 . Let S be an arbitrary stop rule and for $\underline{x} \in S$, there exists a set of sequences $T(\underline{x})$ in \mathcal{X}' whose first n coordinates are the coordinates of \underline{x} . Recall that for each n ,

$$\tau_n = \{\underline{x}_n \in S\}, \tag{5.1}$$

so that if S is closed

$$P_u \left(\bigcup_{n=1}^{\infty} \tau_n \right) = 1.$$

If S is not closed, let

$$E = \{ \chi' - \bigcup_{n=1}^{\infty} \bigcup_{\underline{x} \in \tau_n} T(\underline{x}) \}$$

For each n_1 the number of path segments in τ_n is finite, so suppose that $\{\gamma_{n1}, \gamma_{n2}, \dots, \gamma_{nk}\}$ is a partition of τ_n , and define

$$S_{\pi} = \{\gamma_{11}, \dots, \gamma_{1n_1}, \gamma_{21}, \dots, \gamma_{2n_2}, \dots\} \quad (5.2)$$

to be a stop rule partition S_{π} of the stop rule S .

For purposes of this chapter, the sample space \mathcal{Q} contains E and the elements of S_{π} and the parameter space Θ consists of all pairs of positive integers W and B . Henceforth, we assume that β is fixed.

We now define bounded completeness for Polya sampling for stop rule partitions which are not necessarily closed, nor GMS.

Definition 5.1 The stop rule S is defined to be boundedly complete with respect to Polya sampling if for every bounded real valued function $m(\cdot)$, defined on \mathcal{Q} such that $m(E) = 0$

$$\sum_{\gamma \in \mathcal{Q}} m(\gamma) P_u(\gamma) = 0,$$

for all $W, B \in \Theta$, implies

$$m(\gamma) \equiv 0$$

on \mathcal{Q} .

This definition of bounded completeness is equivalent to requiring that

$$\sum_{\gamma \in S_{\pi}} m(\gamma) P_u(\gamma) = 0$$

for all $W, B \in \Theta$ implies

$$m(\gamma) \equiv 0$$

on S_{π} .

It is useful to associate a set of points, $C(S_{\pi})$, in the plane with the stop rule partition S_{π} . This set $C(S_{\pi})$, constructed as follows, will be called the derived point configuration of S_{π} .

Consider the sets T_n and γ_{nj} of (5.1) and (5.2). For each $\underline{x}_n \in T_n$, define

$$d(\underline{x}_n) = \sum_{i=1}^n x_i$$

and

$$z(\underline{x}_n) = n - d(\underline{x}_n),$$

and for each $\gamma_{nj} \in S_{\pi}$, define

$$d(\gamma) = \min_{\underline{x}_n \in \gamma} d(\underline{x}_n)$$

and

$$z(Y) = \min_{\underline{x}_n \in Y} z(\underline{x}_n).$$

Each of the elements Y of S_π can thus be represented by a point $(d(Y) + z(Y), d(Y)) = (n(Y), d(Y))$, of the derived point configuration $C(S_\pi)$.

Note that the mapping $S_\pi \rightarrow C(S_\pi)$ may be many-to-one, so that S_π is not necessarily characterized by $C(S_\pi)$.

Since $C(S_\pi)$ is a set of points in the (n, d) -plane, a subset $L(S_\pi)$, called the lining of $C(S_\pi)$, may be thought of as a boundary in the sense used by Girshick, Mosteller, and Savage (1946). That is, all points $\alpha \in L(S_\pi)$ may be reached by a sample path from the origin which does not contain another point of $L(S_\pi)$.

Theorem 5.1 Let S_π be a stop rule partition. If (i) the mapping of S_π to $C(S_\pi)$ is one-to-one,

$$(ii) \ L(S_\pi) = C(S_\pi),$$

and (iii) $L(S_\pi)$ is simple, then S is boundedly complete with respect to Polya sampling.

Proof: If $L(S_\pi) = C(S_\pi)$, and $L(S_\pi)$ is simple, it is decomposable, so the points of $L(S_\pi)$ may be associated with a certain sequence of cuts. It is possible to order the sets of S_π by the order in

which the corresponding points of $L(S_\pi)$ are associated with the sequence of cuts. Since the mapping of S_π to $L(S_\pi)$ is one-to-one, there is a unique set γ in S_π for each point α in $L(S_\pi)$.

Suppose that there exists a real-valued bounded function $m(\gamma)$, such that $|m(\gamma)| \leq m^* < \infty$, and

$$\sum_{\gamma \in S_\pi} m(\gamma) P_u(\gamma) = 0 \quad (5.3)$$

for a fixed B and W , $B \in \Theta$. We must show that

$$m(\gamma) = 0$$

for all $\gamma \in S_\pi$.

Suppose γ^* is the first set of S_π , with respect to the ordering imposed on $L(S_\pi)$, for which

$$m(\gamma^*) \neq 0.$$

Let " $\gamma > \gamma^*$ " denote the sets γ which follow γ^* in the ordering of S_π , then (5.3) becomes

$$m(\gamma^*) P_u(\gamma^*) + \sum_{\gamma > \gamma^*} m(\gamma) P_u(\gamma) = 0$$

for all W , $B \in \Theta$, which implies

$$|m(\gamma^*)| \leq \frac{\sum_{\gamma > \gamma^*} |m(\gamma)| P_u(\gamma)}{P_u(\gamma^*)}$$

$$\leq \frac{\frac{m^* \sum_{\gamma > \gamma^*} P_u(\gamma)}{W [d(\gamma^*)]} \cdot \frac{W [d(\gamma^*)]}{N [d(\gamma^*)]}}{P_u(\gamma^*)},$$

for all positive integers W and B.

Since γ^* has at least one element, \underline{x}^* , such that $d(\underline{x}^*) = d(\gamma^*)$, this implies

$$\frac{\frac{W [d(\gamma^*)]}{N [d(\gamma^*)]}}{P_u(\gamma^*)} =$$

$$\frac{1}{\sum_{\substack{\underline{x} \in \gamma^* \\ d(\underline{x}) = d(\gamma^*)}} \frac{B [z(\underline{x}^*)]}{(N + d(\underline{x}^*) \beta)^{[z(\underline{x}^*)]}} + \sum_{\substack{\underline{x} \in \gamma^* \\ d(\underline{x}) \neq d(\gamma^*)}} \frac{(W + d(\underline{x}^*) \beta)^{[d(\underline{x}) - d(\underline{x}^*)]} B [z(\underline{x}^*)]}{(N + d(\underline{x}^*) \beta)^{[d(\underline{x}) - d(\underline{x}^*) + z(\underline{x})]}}.$$

Decomposability and the one-to-one correspondence between elements of S_{Π} and $C(S_{\Pi})$ imply that $d(\gamma) > d(\gamma^*)$, for all $\gamma > \gamma^*$, so

$$\frac{P_u(\gamma)}{\frac{W[d(\gamma^*)]}{N[d(\gamma^*)]}} = \sum_{\underline{x} \in \gamma} \frac{(W+d(\gamma^*)\beta)^{[d(\underline{x})-d(\gamma^*)]} B[z(\underline{x})]}{(N+d(\gamma^*)\beta)^{[d(\underline{x})-d(\gamma^*)+z(\underline{x})]}} .$$

Since (5.3) must hold for all integers W and B , choose a sequence $\{W_n, B_n\}$ such that $N_n \xrightarrow{n \rightarrow \infty} \infty$ and $\frac{W_n}{N_n} \xrightarrow{n \rightarrow \infty} 0$. Then for all $\underline{x} \in \gamma > \gamma^*$,

$$\frac{(W_n+d(\gamma^*)\beta)^{[d(\underline{x})-d(\gamma^*)]} B_n[z(\underline{x})]}{(N_n+d(\gamma^*)\beta)^{[d(\underline{x})-d(\gamma^*)+z(\underline{x})]}} \xrightarrow{n \rightarrow \infty} 0, \quad (5.4)$$

and

$$\begin{aligned} & \frac{1}{\frac{B_n[z(x^*)]}{(N_n+d(\underline{x}^*)\beta)^{[z(\underline{x}^*)]}} + \sum_{\underline{x} \in \gamma^*} \frac{(W_n+d(\underline{x}^*)\beta)^{[d(\underline{x})-d(\underline{x}^*)]} B_n[z(\underline{x})]}{(N_n+d(\underline{x}^*)\beta)^{[d(\underline{x})-d(\underline{x}^*)+z(\underline{x})]}}} \\ & \xrightarrow{n \rightarrow \infty} 1. \end{aligned} \quad (5.5)$$

Therefore $|m(\gamma^*)| \leq 0$ by (5.3), (5.4) and (5.5). In the same manner, it can be shown, for each successive $\gamma \in S_{\Pi}$, that $m(\gamma) = 0$.

qed

The importance of the parameter space enters the proof in (5.4) and (5.5) where we force terms of the type

$$\frac{W_n[d] B_n[z]}{N_n[n]}$$

to zero. Notice that if the parameter space had been $\theta' = \{\beta = 0, \text{ all positive integers } W, B\}$, (5.4) and (5.5) still go to zero, leading to the following

Corollary 5.1a Let S_π be a stop rule partition whose derived point configuration $C(S_\pi)$ satisfies conditions (i), (ii), and (iii) of the theorem, then S_π is boundedly complete with respect to Bernoulli sampling for the parameter space $\theta' = \{p: p = \frac{W}{W+B}, W, B \text{ positive integers}\}$.

Since θ' is a subset in $(0,1)$, the above result easily leads to a sufficient condition for completeness of arbitrary stop rules partitions for Bernoulli sampling situations with the natural parameter space, $\theta'' = (0,1)$.

Corollary 5.1b Let S_π be a stop rule partition whose derived point configuration $C(S_\pi)$ satisfies (i), (ii), and (iii) of the theorem, then S_π is boundedly complete with respect to Bernoulli sampling with parameter space θ'' .

Analogously to Chapter III, we now relate completeness with respect to Bernoulli sampling to completeness with respect to Polya sampling, and conversely.

For the sake of ease of exposition we introduce terms that distinguish between completeness with respect to Polya and Bernoulli sampling procedures.

Definition 5.2 A stop rule partition S_π is called Polya-complete with respect to the class \mathcal{J} of functions defined on S_π , if, for $m \in \mathcal{J}$,

$$\sum_{Y \in S_\pi} m(Y) P_u(Y) = 0$$

for a fixed positive β and all $W, B \in \theta$ implies that

$$m(Y) \equiv 0.$$

Definition 5.3 A stop rule partition S_π is called Bernoulli-complete with respect to the class \mathcal{J} of functions defined on S_π , if, for $m \in \mathcal{J}$,

$$\sum_{Y \in S_\pi} m(Y) P_b(Y) = 0$$

for all $p \in \theta''$ implies that

$$m(Y) \equiv 0.$$

Before proceeding to our utilization of the theorems of Section 2.4, we require another condition on a stop rule partition.

Definition 5.4 A stop rule partition S_{Π} is minimal if $\underline{x}' \in \gamma_1$, $\underline{x}'' \in \gamma_2$, with $\gamma_1 \neq \gamma_2$, implies that $(n', d') \neq (n'', d'')$.

For each set γ in the minimal stop rule partition S_{Π} , let $r(\gamma)$ be the number of distinct values of $\sum_{i=1}^{n(\gamma)} x_i$, and $k_1(\gamma)_1, k_2(\gamma), \dots, k_r(\gamma)$ represent the number of path segments belonging to γ associated with each value, $d_1(\gamma), d_2(\gamma), \dots, d_{r(\gamma)}(\gamma)$.

There are four theorems, two dealing with the Polya-to-Bernoulli direction, and the other two with the converse. They are applications of the four theorems of Section 2.4, which relate arbitrary sums of "Polya-like" and "Bernoulli-like" terms.

Each of the theorems 5.2 to 5.5 allows us to establish a correspondence between Polya-completeness and Bernoulli-completeness under the assumption of minimality for a restricted set of statistics, whose composition is determined by the theorem used.

Bringing to bear Theorem 2.4.1 involves the following construction:

For each (i, j) , $i, j = 0, 1, \dots$, and each function $m(\gamma)$ on a stop rule partition S_{Π} into the reals, define

$$A_{ij}^m = m(\gamma)k_i(\gamma),$$

if there is a path segment in any $\gamma \in S_{\Pi}$ for which $(d_i(\gamma), z_i(\gamma)) = (i, j)$

$$A_{ij}^m = 0,$$

if there is no such path segment in any $\gamma \in S_{\Pi}$. (5.6)

For the purposes of Theorem 5.2, the class \mathcal{S} is the class \mathcal{S}_1 of functions m such that condition (i) of Theorem 2.4.1 holds for $A_{ij} = A_{ij}^m$ for all $W, B \in \theta$, in other words, \mathcal{S}_1 is the class of statistics whose expectation is absolutely summable.

Theorem 5.2 If the minimal stop rule partition S_π is Polya-complete with respect to \mathcal{S}_1 , it is Bernoulli-complete with respect to \mathcal{S}_1 .

Proof: Suppose that $m(Y), Y \in S_\pi$, is such that

$$(1) \sum_{Y \in S_\pi} m(Y) P_b(Y) \text{ converges absolutely for } p \in (0,1),$$

$$(2) \sum_{Y \in S_\pi} m(Y) P_b(Y) = 0 \text{ for } p \in (0,1), \text{ and}$$

$$(3) m \in \mathcal{S}_1.$$

Now substituting for $P_b(Y)$ in the above three conditions, we have that

$$(1') \sum_{Y \in S_\pi} m(Y) \sum_{i=1}^{r(Y)} k_i(Y) p^{d_i(Y)} q^{z_i(Y)}$$

converges absolutely by the positivity of $k_i(Y)$, for all $Y \in S_\pi$, so that we may rewrite it as

$$\sum_{\gamma \in S_{\pi}} \sum_{i=1}^{r(\gamma)} m(\gamma) k_i(\gamma) p^{d_i(\gamma)} q^{z_i(\gamma)},$$

(2') in view of (1'), (2) becomes

$$\sum_{\gamma \in S_{\pi}} \sum_{i=1}^{r(\gamma)} m(\gamma) k_i(\gamma) p^{d_i(\gamma)} q^{z_i(\gamma)} = 0$$

for all $p \in (0,1)$, and

$$(3') \quad m \in \mathcal{M}_1.$$

From the minimality of S_{π} , we can use (5.6) to associate a constant A_{ij}^m with each pair (γ, i) , $\gamma \in S_{\pi}$ and $i = 1, \dots, r(\gamma)$, so that the three conditions can now be rewritten as

$$(1'') \quad \sum_i \sum_j A_{ij}^m p^i q^j$$

absolutely convergent for $p \in (0,1)$,

$$(2'') \quad \sum_i \sum_j A_{ij}^m p^i q^j = 0,$$

for $p \in (0,1)$.

$$(3'') \quad A_{ij}^m \text{ satisfy condition (i) of Theorem 2.4.1 for each}$$

$w, B \in \theta$.

In view of Theorem 2.4.1,

$$\sum_i \sum_j A_{ij}^m \frac{W^{[i]} B^{[j]}}{N^{[i+j]}} = 0$$

for all $W, B \in \Theta$, and in view of condition (i) of the theorem.

$$\sum_{\gamma \in S_{\pi}} m(\gamma) \sum_{i=1}^{r(\gamma)} k_i(\gamma) \frac{W^{[d_i(\gamma)]} B^{[z_i(\gamma)]}}{N^{[n(\gamma)]}} = 0.$$

Since S_{π} is Polya-complete with respect to \mathcal{L}_1 ,

$$m(\gamma) \equiv 0.$$

qed

The proofs of the next two theorems are similar to that of Theorem 5.2, therefore the theorems will be stated without proof.

For Theorem 5.3, the class \mathcal{L} is the class \mathcal{L}_2 of functions m such that $A_{ij} = A_{ij}^m$ satisfies the conditions of Theorem 2.4.2 for all $W, B \in \Theta$.

Theorem 5.3 If the minimal stop rule partition S_{π} is Polya-complete with respect to \mathcal{L}_2 , it is Bernoulli-complete with respect to \mathcal{L}_2 .

In the case of Theorem 5.4, the class \mathcal{J} is the set \mathcal{J}_3 of functions m for which $A_{ij} = A_{ij}^m$ satisfy conditions (i) and (iii) of Theorem 2.4.3 for every $p^* \in (0,1)$.

Theorem 5.4 If the minimal stop rule partition S_π is Bernoulli-complete with respect to \mathcal{J}_3 it is Polya-complete with respect to \mathcal{J}_3 .

Since Theorem 5.5 is relied upon in Chapter VI to relate the existence of a UMVU estimate of p under Polya sampling to its existence under Bernoulli sampling, we provide a detailed proof.

The class is now the class \mathcal{J}_4 of bounded functions m for which $A_{ij} = A_{ij}^m$ satisfy condition (i) of Theorem 2.4.4 for all $p^* \in (0,1)$.

Theorem 5.5 If the minimal stop rule partition S_π is Bernoulli-complete with respect to \mathcal{J}_4 it is Polya-complete with respect to \mathcal{J}_4 .

Proof: Suppose that $m(\gamma), \gamma \in S_\pi$ is such that

$$(1) \quad \sum_{\gamma \in S_\pi} m(\gamma) P_u(\gamma) \text{ converges absolutely for all } W, B \in \theta,$$

$$(2) \quad \sum_{\gamma \in S_\pi} m(\gamma) P_u(\gamma) = 0$$

for all $W, B \in \theta$, and

$$(3) \quad m \in \mathcal{J}_4.$$

Substituting for $P_u(\gamma)$ in (1) - (3) yields

$$(1') \quad \sum_{\gamma \in S_{\Pi}} m(\gamma) \sum_{i=1}^{r(\gamma)} k_i(\gamma) \frac{W^{[d_i(\gamma)]} B^{[z_i(\gamma)]}}{N^{[n(\gamma)]}}$$

converges absolutely for all $W, B \in \theta$, so that it may be rewritten as

$$\sum_{\gamma \in S_{\Pi}} \sum_{i=1}^{r(\gamma)} m(\gamma) k_i(\gamma) \frac{W^{[d_i(\gamma)]} B^{[z_i(\gamma)]}}{N^{[n(\gamma)]}},$$

(2') in view of (1'),

$$\sum_{\gamma \in S_{\Pi}} \sum_{i=1}^{r(\gamma)} m(\gamma) k_i(\gamma) \frac{W^{[d_i(\gamma)]} B^{[z_i(\gamma)]}}{N^{[n(\gamma)]}} = 0$$

for all $W, B \in \theta$, and

(3') $m \in \mathcal{S}_4$.

Utilizing (5.6) by the minimality of S_{Π} , we associate a constant A_{ij}^m with each pair (γ, i) , $\gamma \in S_{\Pi}$ and $i=1, \dots, r(\gamma)$ so that (1') and (2') can be rewritten as

$$(1'') \quad \sum_i \sum_j A_{ij}^m \frac{W^{[i]}}{N^{[i+j]}}$$

converges absolutely for all $W, B \in \theta$,

$$(2'') \quad \sum_i \sum_j A_{ij}^m \frac{W^{[i]} B^{[j]}}{N^{[i+j]}} = 0$$

for all $W, B \in \mathcal{D}$.

Fix $p^* \in (0,1)$. Now $(2'')$ holds on any sequence $\{W_u, B_u\}_{u=1}^{\infty}$ for which $\frac{W_u}{N_u} \xrightarrow{u} p^*$. In view of $(1'')$, let $(i_v, j_v)_{v=1}^{\infty}$ be the ordering for which condition (i) of Theorem 2.4.4 holds. Condition (iii) of the theorem is assured by the boundedness of m .

Now applying Theorem 2.4.4,

$$\sum_v A_v^m \frac{W_u^{[i_v]} B_u^{[j_v]}}{N_u^{[i_v+j_v]}} = 0,$$

for all u , implies that

$$\sum_v A_v^m p^{*i_v} q^{*j_v} = 0.$$

The boundedness of m assures the absolute convergence of

$\sum_i \sum_j A_{ij}^m p^{*i} q^{*j}$, so that

$$\sum_i \sum_j A_{ij}^m p^{*i} q^{*j} = 0. \quad (5.7)$$

Substituting for A_{ij}^m , and by absolute convergence, we may rewrite

(5.7) as

$$\sum_{\gamma \in S_{\pi}} m(\gamma) \sum_{i=1}^{r(\gamma)} k_i(\gamma) p^{*di(r)} q^{*zi(r)} = 0.$$

Since S_{π} is Bernoulli-complete with respect to \mathcal{S}_4 ,

$$m(\gamma) \equiv 0.$$

VI. SEQUENTIAL UMVU ESTIMATION

6.1. Introduction

In Section 6.2, we identify conditions for the existence of a UMVU estimate of p , under Polya sampling, for stop rule partitions which are not necessarily of the GMS type. Part of this section involves a different emphasis of the point configuration $C(S_{\pi})$ and lining $L(S_{\pi})$. $C(S_{\pi})$ and $L(S_{\pi})$ originally were introduced in Chapter IV as geometrically constructive devices, and in Chapter V, under the assumption $C(S_{\pi}) = L(S_{\pi})$, served to relate the essentially geometric notions of simplicity and decomposability to the bounded completeness of general, not necessarily GMS stop rule partitions. In Theorem 6.2.1 $C(S_{\pi})$ and $L(S_{\pi})$ perform a similar function under the additional conditions of minimality and sufficiency. However, in Example 6.2.1, Lemmas 6.2.2, 6.2.3, and 6.2.4, $L(S_{\pi})$ is now interpreted as a GMS stop rule partition in its own right, certain of whose inferential properties are related to the corresponding properties of S_{π} .

We also show that a finite stop rule partition which satisfies the conditions for the existence of UMVU estimate of p given by Theorem 6.2.1, is a GMS stop rule partition.

In Section 6.3, we relate questions concerning UMVU estimates of p in the Bernoulli case to the corresponding questions for Polya sampling, including questions of existence of the UMVU estimates for both sampling procedures, using the relationship between the

respective minimum variances.

Section 6.4 presents a sequential generalization of the finite population correction to Polya sampling.

6.2. Sequential Polya UMVU Estimates

We begin with a formulation of sufficiency suitable to sequential, not necessarily GMS, stop rule partitions.

Definition 6.2.1 A stop rule partition S_π is sufficient if the conditional probability of a path segment, given the partition set in S_π to which it belongs, is independent of the parameters.

Lemma 6.2.1 A stop rule partition S_π is (a) sufficient with respect to Polya sampling, if and only if (b) it is sufficient with respect to Bernoulli sampling, if and only if (c) for all $\gamma \in S_\pi$ all path segments contained in γ have the same value of n and

$$\sum_{i=1}^n x_i.$$

Proof: (a) \Rightarrow (b) If S_π is sufficient for Polya sampling (resp., Bernoulli sampling), then for each $\gamma \in S_\pi$ and each path segment $\underline{x}' \in \gamma$, the conditional probability of \underline{x}' given γ ,

$$P_u(\underline{x}' | \gamma) = \frac{\frac{W}{N} \frac{[d(\underline{x}')] [z(\underline{x}')] }{[n(\gamma)]}}{\sum_{\underline{x} \in \gamma} \frac{W}{N} \frac{[d(\underline{x})] [z(\underline{x})]}{[n(\gamma)]}} = K(\gamma), \quad (6.2.1)$$

for all W and B ,

(resp.,

$$P_b(\underline{x}' | \gamma) = \frac{\prod_{p \in \gamma} d(\underline{x}') \prod_{q \in \gamma} z(\underline{x}')}{\prod_{p \in \gamma} d(\underline{x}) \prod_{q \in \gamma} z(\underline{x})} = K(\gamma) \quad (6.2.2)$$

for all p),

where $K(\gamma)$ is a constant and $d(\underline{x})$ and $z(\underline{x})$ are the values of

$$\sum_{i=1}^n x_i \text{ and } n - \sum_{i=1}^n x_i \text{ corresponding to } \underline{x}.$$

Since the number of path segments in each set γ is finite, we can apply Corollary 2.4.4b, so that (6.2.1) holds if and only if (6.2.2) holds.

(b) \nRightarrow (c) If S is sufficient in the Bernoulli context, then

$$\begin{aligned} P_b(\underline{x}' | \gamma) &= K(\gamma) = \frac{\prod_{p \in \gamma} d(\underline{x}') \prod_{q \in \gamma} z(\underline{x}')}{\prod_{p \in \gamma} d(\underline{x}) \prod_{q \in \gamma} z(\underline{x})} \\ &= \frac{1}{1 + \sum_{p \in \gamma} (d(\underline{x}) - d(\underline{x}')) \prod_{q \in \gamma} (z(\underline{x}) - z(\underline{x}'))} \end{aligned}$$

for all p .

This implies that

$$\sum_{\underline{x} \in \gamma} p^{(d(\underline{x})-d(\underline{x}'))} q^{(z(\underline{x})-z(\underline{x}'))} = \frac{1}{K(\gamma)} - 1,$$

for all p ,

which in turn implies that

$$d(\underline{x}) = d(\underline{x}')$$

and

$$z(\underline{x}) = z(\underline{x}')$$

for all $\underline{x} \in \gamma$.

(c) \supseteq (b) If for all $\underline{x} \in \gamma$, $d(\underline{x}) = d$ and $z(\underline{x}) = z$, then

$$P_b(\underline{x}|\gamma) = \frac{p^d q^z}{k(\gamma) p^d q^z} = \frac{1}{k(\gamma)},$$

where $k(\gamma)$ is the number of path segments in γ .

qed

Remark 6.2.1 In view of part of the proof of $((a) \not\Rightarrow (b))$, a stop rule partition S_π may be identified as sufficient, without specifying whether the sampling procedure is a Polya or Bernoulli one.

Remark 6.2.2 In view of the lemma, there is, for each γ of a sufficient S_π , a unique associated value of $\sum_{i=1}^{n_\gamma} x_i$, ($n_\gamma = n(\gamma)$), henceforth denoted by d_γ . It follows that the point of $C(S_\pi)$ corresponding to a given γ is now simply

$$(d(\gamma) + z(\gamma), d(\gamma)) = (n_\gamma, d_\gamma) \equiv (n, d)_\gamma.$$

Remark 6.2.3 The properties of S_π , minimality (defined in Chapter V) and sufficiency, refer simultaneously to both sampling procedures.

Remark 6.2.4 When S_π is minimal and sufficient, there is a unique $\alpha \in C(S_\pi)$, namely $(n, d)_\gamma \equiv \alpha(\gamma)$, corresponding to any $\gamma \in S_\pi$, and a unique $\gamma \in S_\pi$, say $\gamma(\alpha)$, with $(n, d)_{\gamma(\alpha)} = (n, d)$, corresponding to any given $\alpha = (n, d) \in C(S_\pi)$.

We are now in a position to state conditions under which there exists a UMVU estimate for p in the context of Polya sampling.

Theorem 6.2.1 extends the work of Girshick, Mosteller, Savage (1946) and Wolfowitz (1946) in two directions: to Polya sampling and to stop rule partitions which are not necessarily GMS. Two immediate consequences of the theorem are that the estimate given by

the theorem is the UMVU estimate for p (i) under Polya sampling for a GMS stop rule partition, and (ii) under Bernoulli sampling for certain non-GMS stop rule partitions.

Theorem 6.2.1 If the stop rule partition S_{π} is (i) closed and (ii) minimal sufficient, (iii) $L(S_{\pi}) = C(S_{\pi})$, and (iv) $L(S_{\pi})$ is simple, then there exists a bounded-UMVU estimate of p , under Polya sampling, given by

$$t(\gamma) = \frac{k^*(\gamma)}{k(\gamma)}, \quad (6.2.3)$$

$\gamma \in S_{\pi}$, where $k(\gamma)$ is the number of path segments belonging to γ , and $k^*(\gamma)$ is the number of path segments, among the $k(\gamma)$ path segments belonging to γ , for which $x_1 = 1$.

Proof: Define a statistic u on the path segments \underline{x} of S as follows:

$$u(\underline{x}) = \begin{cases} 1 & \text{if } x_1 = 1 \\ 0 & \text{if } x_1 = 0. \end{cases}$$

Since S is closed,

$$\begin{aligned} E_u(u(\underline{x})) &= 1 \cdot \sum_{\substack{\underline{x} \in S \\ \exists x_1=1}} P_u(\underline{x}) + 0 \cdot \sum_{\substack{\underline{x} \in S \\ \exists x_1=0}} P_u(\underline{x}) \\ &= p, \end{aligned}$$

so $u(\cdot)$ is unbiased for p .

By conditioning u on the sufficient partition S_{π} , we produce a statistic on S_{π} :

$$\begin{aligned}
 t(\gamma) &= E_u (u(\underline{x}) | \gamma) \\
 &= 1 \cdot P_u (u(\underline{x}) = 1 | \gamma) + 0 \cdot P_u (u(\underline{x}) = 0 | \gamma) \\
 &= P_u (u(\underline{x}) = 1 | \gamma) \\
 &= \frac{P_u (\underline{x} \in \gamma \text{ and } u(\underline{x}) = 1)}{P_u (\underline{x} \in \gamma)} \\
 &= \frac{k^*(\gamma)}{k(\gamma)} ,
 \end{aligned}$$

which is unbiased for p and bounded by one.

In view of Remark 6.2.1, minimality and sufficiency together imply the one-to-one correspondence between S_{π} and $C(S_{\pi})$, so that conditions (ii), (iii), and (iv) ensures the bounded Polya-completeness of S_{π} by Theorem 5.1. Therefore, t is the only bounded unbiased estimate for p on S_{π} .

Let s be any other unbiased estimate of p on S . We may condition s on S_π , as before, to get an unbiased, bounded estimate t' on S_π . By Jensen's Inequality for conditional expectations (see Theorem 9.1.4, page 281 of Chung [1968]),

$$E_u(s^2) \geq E_u(t'^2),$$

so that

$$\text{Var}_u(s) \geq \text{Var}_u(t').$$

The bounded completeness of S_π implies that $t' = t$; therefore

$$\text{Var}_u(s) \geq \text{Var}_u(t),$$

so that t is the bounded-UMVU estimate of p under Polya sampling.

qed

One may well be concerned that the conditions imposed on S_π by Theorem 6.2.1 may restrict the class of stop rules to which it applies to only those which are GMS. This is not the case.

Example 6.2.1 Suppose we have a stop rule S whose members are as follows:

(1) (1,1,1,0)

(2) (1,1,1,1)

- (3) for all $n \geq 4$, all path segments (x_1, \dots, x_n) with $\sum_{i=1}^n x_i = n-4$, except those for which $(x_1, x_2, x_3, x_4) = (1, 1, 1, 0)$ or $(1, 1, 1, 1)$.

We define the stop rule partition S_π to be

$$\gamma_1 = (1, 1, 1, 0)$$

$$\gamma_2 = (1, 1, 1, 1)$$

$$\gamma_{n-1} = \left\{ (x_1, \dots, x_n) : \sum_{i=1}^n x_i = n-4 \quad \text{and} \right. \\ \left. (x_1, x_2, x_3, x_4) \neq (1, 1, 1, 0) \quad \text{or} \right. \\ \left. (x_1, x_2, x_3, x_4) \neq (1, 1, 1, 1) \right\},$$

for $n \geq 4$.

Now S_{π_1} is minimal and sufficient since each γ contains all path segments of S which have the same value of n and $\sum_{i=1}^n x_i$. It may be seen that $L(S_{\pi_1}) = C(S_{\pi_1})$ and $L(S_{\pi_1})$ is simple; S_{π_1} , in view of Theorem 3.1, is closed since the number of values of

$\sum_{i=1}^n x_i$ for the non-terminated path segments at each n is three.

By Theorem 6.2.1, S_{π_1} is an infinite stop rule partition which admits a UMVU estimate of p , but is clearly not GMS.

It should be noted, however, that if S_π is a finite stop rule partition and satisfies Theorem 6.2.1 then S_π is GMS.

Lemma 6.2.2 A finite, closed, minimally sufficient stop rule partition S_π for which $L(S_\pi) = C(S_\pi)$ and $L(S_\pi)$ is simple is a GMS stop rule partition.

Proof: Assume that S_π is not a GMS stop rule partition. Since S_π is minimally sufficient, for each $\gamma_i \in S_\pi$ there exists a unique point $\alpha_i \in L(S_\pi)$ such that $(n, x)_{\gamma_i} = (n, x)_{\alpha_i}$. Also, since S_π is not GMS, there exists at least one i such that $k(\gamma_i) \neq k(\alpha_i)$. Let γ_j and α_j be the set of S_π and corresponding point of $L(S_\pi)$ with the smallest value of n for which this occurs.

Claim: $k(\alpha_j) > k(\gamma_j)$. Assume not; then γ_j stops more paths than α_j . Since γ_i and α_i , $i=1, \dots, j-1$ stop the same paths, then γ_j must stop at least one path that had been previously stopped. This contradicts the definition of the stop sets $\{\tau_n\}$, and hence the definition of a stop rule.

Without loss of generality, let the path segment $\underline{x} = (1, 0, 1, 1, 0)$ be stopped by $\alpha_j = (5, 3)$, but not by γ_j . Since S_π is finite, there exists a number n' such that

$$n' = \max_i (n_{\alpha_i}) = \max_i (n_{\gamma_i}).$$

Again, assume, without loss of generality, that $n' = 10$. Consider the two eleven-tuples $\underline{x}' = (1, 0, 1, 1, 0, 1, 1, 1, 1, 1, 1)$ and $\underline{x}'' = (1, 0, 1, 1, 0, 0, 0, 0, 0, 0, 0)$. Since S_π is closed, paths must be terminated with probability one, but $P_u(\underline{x}') > 0$ and $P_u(\underline{x}'') > 0$. Therefore, the path segments represented by \underline{x}' and \underline{x}'' must have been

stopped on or before trial 10. There must now exist two sets of S_π , γ_1' and γ_2' , which contain a truncated version of \underline{x}' and \underline{x}'' , respectively, and associated points of $L(S_\pi)$, $\alpha_1' = (n_1', n_1' - 2)$ and $\alpha_2' = (n_2', 3)$, where $6 < n_1', n_2' \leq 10$.

Since $L(S_\pi)$ is simple, it is decomposable by Theorem 5.2; therefore, α_j must lie on one of the cuts, $y = x-2$ or $y = 3$. Assume α_j lies on the cut $y = x-2$; but α_1' also lies on $y = x-2$, which violates decomposability. The same contradictory situation exists if α_j lies on the cut $y = 3$.

qed

At this point we change our emphasis of $L(S_\pi)$, as was mentioned in Section 6.1, and now interpret it as a GMS stop rule partition in its own right, that is, as a "boundary." We relate the Polya-completeness of S_π to the Polya-completeness of the stop rule defined by the boundary $L(S_\pi)$, and conversely, without resorting to simplicity and being tied to bounded completeness.

Remark 6.2.5 When S_π is minimal and sufficient and $C(S_\pi) = L(S_\pi)$, then the one-to-one map $(\gamma(\alpha), \alpha(\gamma))$ of Remark 6.2.4 of course involves S_π and $L(S_\pi)$. Moreover, in view of $L(S_\pi) = C(S_\pi)$, all the sets γ of S_π , and points α of $L(S_\pi)$ involved in this map have associated path counts greater than zero.

The following two lemmas concern the Polya-completeness of S_{π} and the Polya-completeness of $L(S_{\pi})$; only the second meshes into our development of UMVU estimates, but the first is natural.

Lemma 6.2.1 If S_{π} is minimal and sufficient and $L(S_{\pi}) = C(S_{\pi})$, then S_{π} is Polya-complete if and only if $L(S_{\pi})$ is Polya-complete.

Proof: Consider first a sequence of constants $\{g_i\}_{i=1}^{\infty}$ such that

$$\sum_i g_i P_u(\alpha_i) = 0,$$

$\alpha_i \in L(S_{\pi})$. By the assumptions of the theorem and Remarks 6.2.4 and 6.2.5, we may write

$$P_u(\alpha_i) = \frac{k(\alpha_i)}{k(\gamma_i)} P_u(\gamma_i).$$

Then

$$\sum_i g_i P_u(\alpha_i) = \sum_i g_i \frac{k(\alpha_i)}{k(\gamma_i)} P_u(\gamma_i),$$

and in view of Remark 6.5.1, the Polya-completeness of S_{π} implies that

$$g_i = 0$$

for all i . The proof in the other direction interchanges $P_u(\alpha)$ and $P_u(\gamma)$. qed

The next lemma will be stated without proof; its proof is essentially the same as that of the previous lemma.

Lemma 6.2.2 Denote by K_i the ratios $k(\alpha_i)/k(\gamma(\alpha_i))$ appearing in the proof of Lemma 6.2.1. Suppose that S_π is minimal and sufficient, $C(S_\pi) = L(S_\pi)$, and that there are constants Q_1 and Q_2 such that

$$0 < Q_1 < K_i < Q_2 < +\infty.$$

Then S_π is boundedly Polya-complete if and only if $L(S_\pi)$ is boundedly Polya-complete.

6.3. Polya UMVU Estimates and Bernoulli UMVU Estimates

The existences of UMVU estimates under Polya and Bernoulli sampling are related in the theorems of this section.

For ease of exposition, we define terms for unbiasedness under Polya and Bernoulli sampling.

Definition 6.3.1 A statistic s is Polya-unbiased for p if $E_u(s) = p$, for all W , B , and β .

Definition 6.3.2 A statistic s is Bernoulli-unbiased for p if $E_b(s) = p$, for all $p \in (0,1)$.

Theorem 6.3.1 If the stop rule partition S_{π} is closed under Polya sampling, minimal, sufficient, and Bernoulli-complete with respect to \mathcal{S}_4 and t (see (6.2.3)) belongs to \mathcal{S}_4 , then t is the UMVU estimate of p with respect to \mathcal{S}_4 under Polya sampling.

Proof: Since S_{π} is minimal and Bernoulli-complete with respect to \mathcal{S}_4 , by Theorem 5.5, S_{π} is also Polya-complete with respect to \mathcal{S}_4 .

By sufficiency and closure under Polya sampling, t is Polya-unbiased for p .

Let s be a Polya-unbiased estimate of p . We may condition s on S_{π} to obtain an unbiased statistic t' , such that

$$\text{Var}_u(s) \geq \text{Var}_u(t')$$

by Jensen's Inequality (Theorem 9.1.4, page 281 of Chung [1968]).

If $t' \in \mathcal{S}_4$, then by the Polya-completeness with respect to \mathcal{S}_4 of S_{π} , $t' = t$, so that

$$\text{Var}_u(s) \geq \text{Var}_u(t).$$

Hence t is UMVU with respect to \mathcal{S}_4 under Polya sampling.

Theorem 6.3.2 If the stop rule partition S_{π} is closed under Bernoulli sampling, minimal, sufficient, and Polya-complete with respect to \mathcal{S}_1 then t (see (6.2.3)) is the UMVU estimate of p with respect to \mathcal{S}_1 , under Bernoulli sampling.

Proof: Since S_{π} is closed with respect to Bernoulli sampling and sufficient, t is Bernoulli-unbiased for p . Also, since t is bounded, $E(|t|)$ is finite, so that $t \in \mathcal{J}_1$.

By Theorem 5.2, minimality and Polya-completeness imply the Bernoulli completeness of S_{π} .

Let s be any Bernoulli-unbiased estimate of p . Conditioning s on S_{π} yields an unbiased statistic t' on S_{π} for which, by Jensen's Inequality (Theorem 9.1.4, page 281 of Chung [1968]),

$$\text{Var}_b(s) \geq \text{Var}_b(t').$$

Since S_{π} is Bernoulli-complete, $t' = t$, and

$$\text{Var}_b(s) \geq \text{Var}_b(t),$$

so that t is the UMVU estimate of p under Bernoulli sampling.

Theorem 6.3.3 Let \mathcal{J}_5 be the class of Bernoulli-unbiased statistics m which have finite Bernoulli second moments. If t (see (6.2.3)) is a member of \mathcal{J}_5 and has uniformly smallest Bernoulli variance among the statistics in \mathcal{J}_5 , then t is the UMVU estimate of p with respect to \mathcal{J}_5 under Polya sampling.

Proof: Since s is Bernoulli-unbiased, it is Polya-unbiased by Theorem 2.4.1.

Let m be any other statistic of \mathcal{J}_5 , then

$$E_b(m^2) \geq E_b(t^2),$$

for all $p \in (0,1)$. Since $m \in \mathcal{J}_5$, by Lemma 2.4.1,

$$E_u(m^2) \geq E_u(t^2)$$

for all $W, B \in \theta$, so that

$$\text{Var}_u(m) \geq \text{Var}_u(t)$$

for all $W, B \in \theta$, therefore, t is the UMVU estimate of p under Polya sampling.

6.4. Sequential Polya Correction

In Section 2.3, we introduced a fixed-sample-size generalization, for Polya sampling, of the well-known finite population correction.

It is possible to generalize the finite population correction in another sense--namely from fixed-sample-size sampling to sequential sampling.

Suppose that Mary has in hand a certain statistic s that she knows to be unbiased for p under Bernoulli sampling with respect to a certain stop rule S and has computed the variance of s under S as a polynomial $\text{Var}_b(s)$ in p . Suppose also that Mary tells Jane just (1) that s is Bernoulli unbiased for p under S and (2) the polynomial $\text{Var}_b(s)$, but tells her nothing else--in particular, shows her neither s nor S . It is possible for Jane, using only the information available to her, to compute the variance $\text{Var}_k(s)$ of s under

S and hypergeometric sampling from a dichotomized finite population of size $W + B \equiv N$, assuming only that W and B are large enough so that S can be carried out without truncation.

Jane's simple general recipe for computing $\text{Var}_h(s)$ from $\text{Var}_b(s)$ is the "sequential finite population correction" formulated by David and Olkin (1956).

There is an analogous simple recipe for computing $\text{Var}_u(s)$ from $\text{Var}_b(s)$, and it will be called the sequential Polya correction and its form is given in Theorem 6.4.1.

Theorem 6.4.1 Consider a minimal stop rule partition S_π and a Bernoulli-unbiased statistic s defined on S_π . If the variance of s under Bernoulli sampling is given by

$$\sum_{\gamma \in S_\pi} \sum_{i=1}^{r(\gamma)} A(\gamma, i) p_i^{d_i(\gamma)} q^{(n(\gamma)-d_i(\gamma))} \quad (6.4.1)$$

and

$$\sum_{\gamma \in S_\pi} s(\gamma) \sum_{i=1}^{r(\gamma)} k_i(\gamma) \frac{\frac{[d_i(\gamma)] [n(\gamma)-d_i(\gamma)]}{B}}{\frac{[n(\gamma)]}{N}} \quad (6.4.2)$$

and

$$\sum_{\gamma \in S_\pi} s^2(\gamma) \sum_{i=1}^{r(\gamma)} k_i(\gamma) \frac{\frac{[d_i(\gamma)] [n(\gamma)-d_i(\gamma)]}{B}}{\frac{[n(\gamma)]}{N}} \quad (6.4.3)$$

both satisfy condition (i) of Theorem 2.4.1, then the variance of s with respect to Polya sampling is given by

$$\sum_{\gamma \in S_{\pi}} \sum_{i=1}^{r(\gamma)} A(\gamma, i) \frac{\binom{d_i(\gamma)}{W} \binom{n(\gamma)-d_i(\gamma)}{B}}{\binom{n(\gamma)}{N}} + \frac{\binom{2}{W}}{\binom{2}{N}} - p^2.$$

Proof: Since s is Bernoulli-unbiased, and in view of (6.4.2), s is Polya-unbiased by Theorem 2.4.1.

Now

$$\begin{aligned} \text{Var}_b(s) &= \sum_{\gamma \in S_{\pi}} \sum_{i=1}^{r(\gamma)} A(\gamma, i) p^{d_i(\gamma)} q^{(n(\gamma)-d_i(\gamma))} \\ &= E_b(s^2) - [E_b(s)]^2 \\ &= \sum_{\gamma \in S_{\pi}} s(\gamma)^2 \sum_{i=1}^{r(\gamma)} k_i(\gamma) p^{d_i(\gamma)} q^{(n(\gamma)-d_i(\gamma))} - p^2; \end{aligned}$$

Rewriting $E_b(s^2)$ and $\text{Var}_b(s)$ by the minimality of S_{π} , and utilizing (6.6), we have

$$E_b(s^2) = \sum_{(i,j)} C_{ij}^s p^i q^j = \sum_{(i,j)} A_{ij}^s p^i q^j + p^2 = \text{Var}_b(s) + p^2.$$

Since (6.4.2) and (6.4.1) satisfy condition (i) of Theorem 2.4.1, so also do

$$\sum_{(i,j)} A_{ij}^s \frac{W \frac{[i] [j]}{[i+j]}}{N}$$

and

$$\sum_{(i,j)} C_{ij}^s \frac{W \frac{[i] [j]}{[i+j]}}{N},$$

so, by Theorem 2.4.1,

$$\sum_{(i,j)} C_{ij}^s \frac{W \frac{[i] [j]}{[i+j]}}{N} = \sum_{(i,j)} A_{ij}^s \frac{W \frac{[i] [j]}{[i+j]}}{N} + \frac{W \frac{[2]}{[2]}}{N}. \quad (6.4.4)$$

Rewriting (6.4.4) by the absolute convergence of (6.4.2) and (6.4.3),

we have

$$\text{Var}_u(s) = E_u(s^2) - p^2$$

$$= \sum_{\gamma \in S_{\pi}} r(\gamma) \sum_{i=1} A(\gamma, i) \frac{W \frac{[d_i(\gamma)] [n(\gamma) - d_i(\gamma)]}{[n(\gamma)]}}{N} + \frac{W \frac{[2]}{[2]}}{N} - p^2.$$

qed

VII. BIBLIOGRAPHY

- Anderson, T. W. 1958. An introduction to multivariate statistical analysis. John Wiley and Sons, Inc., New York. 374 pp.
- Audley, R. J., and A. R. Jonckheere. 1956. The statistical analysis of the learning process. The British Journal of Statistical Psychology 9: 87-94.
- Bahadur, R. R. 1954. Sufficiency and statistical decision functions. Ann. Math. Stat. 25: 423-462.
- Buck, R. C. 1965. Advanced calculus. 2nd ed. McGraw-Hill Book Co., New York. 527 pp.
- Bush, R. R., and Frederick Mosteller. 1953. A stochastic model with applications to learning. Ann. Math. Stat. 24: 559-585.
- Chung, Kai Lai. 1968. A course in probability theory. Harcourt, Brace and World, Inc., New York. 331 pp.
- Cochran, W. G. 1963. Sampling techniques. 2nd ed. John Wiley and Sons, Inc., New York. 413 pp.
- David, H. T., and Ingram Olkin. 1956. Sequential estimation from a finite population. Technical Report No. 5, OOR Project No. 1599. University of Chicago, 15 pp.
- DeGroot, M. H. 1959. Unbiased sequential estimation for binomial populations. Ann. Math. Stat. 30: 81-101.
- Dubins, L. E., and David Freedman. 1965. A sharper form of the Borel-Cantelli lemma on the strong law. Ann. Math. Stat. 36: 800-807.
- Feller, William. 1968. An introduction to probability theory and its applications, Volume I. 3rd ed. Wiley and Sons, Inc., New York. 509 pp.
- Feller, William. 1971. An introduction to probability theory and its applications, Volume II. 2nd ed. John Wiley and Sons, Inc., New York. 669 pp.
- Freedman, D. A. 1965. Bernard Friedman's urn. Ann. Math. Stat. 36: 956-970.

- Girshick, M. A., Frederick Mosteller, and L. J. Savage. 1946. Unbiased estimates for certain binomial sampling problems with applications. *Ann. Math. Stat.* 17: 13-23.
- Johnson, N. L., and Samuel Kotz. 1969. *Discrete distributions*. Houghton Mifflin Co., Boston. 328 pp.
- Karlin, S. 1974. Inequalities for symmetric sampling plans I. *Annals of Statistics* 2: 1065-1094.
- Klotz, Jerome. 1973. Statistical inference in Bernoulli trials with dependence. *Annals of Statistics* 1: 373-379.
- Lehmann, E. L. 1949-1950. Notes on theory of estimation. Mimeo. University of California, Berkeley, California. 170 pp.
- Lehmann, E. L., and C. Stein. 1950. Completeness in the sequential case. *Ann. Math. Stat.* 21: 376-385.
- Plackett, R. L. 1948. Boundaries of minimum size in binomial sampling. *Ann. Math. Stat.* 19: 575-580.
- Robbins, Herbert. 1970. Statistical methods related to the law of the iterated logarithm. *Ann. Math. Stat.* 41: 1397-1409.
- Royden, H. L. 1968. *Real analysis*. 2nd ed. The Macmillan Co., Collier-Macmillan Ltd., London. 349 pp.
- Savage, L. J. 1947. A uniqueness theorem for unbiased sequential binomial estimation. *Ann. Math. Stat.* 18: 295-297.
- Sobel, Milton. 1953. An essentially complete class of decision functions for certain standard sequential problems. *Ann. Math. Stat.* 24: 319-337.
- Wald, A. 1947. *Sequential analysis*. John Wiley and Sons, Inc., New York. 212 pp.
- Wolfowitz, J. 1946. On sequential binomial estimation. *Ann. Math. Stat.* 17: 489-493.
- Wolfowitz, J. 1947. Consistency of sequential binomial estimates. *Ann. Math. Stat.* 18: 131-135.

VIII. ACKNOWLEDGMENTS

I would like to thank my two major professors H. T. David and Barry Arnold for introducing me to the topics of Polya urns and sequential estimation; they provided important guidance and valuable suggestions and criticisms during all phases of the research.

I also wish to express my gratitude to Steve Doctor for his support and understanding throughout my degree program.